Forensic Science International 364 (2024) 112244

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Examiner consistency in perceptions of fingerprint minutia rarity☆

Adele Quigley-McBride [a,*,1], Heidi Eldridge [b,2], Brett Gardner [c,3]

[a] Department of Psychology, Simon Fraser University, Canada
[b] Department of Forensic Science, The George Washington University, USA
[c] Institute of Law, Psychiatry, and Public Policy, University of Virginia, USA

ARTICLE INFO

Keywords:
Latent print examination
Minutiae
Base rates
Perceptions
Friction ridge examination

ABSTRACT

Friction ridge examiners (FREs) identify distinctive features (minutiae) in fingerprints and consider how rare these observed minutiae are in their decisions about both the value of a fingerprint and whether there is enough correspondence between two fingerprints to support an "identification" or "exclusion" decision. But subjective perceptions about the frequency of events and features tend to be inconsistent and dynamic, which means that variable perceptions of minutia frequency may contribute to inconsistencies in FREs' opinions about fingerprint evidence. We surveyed expert FREs at two time points ($N_{Time\ 1} = 132$; $N_{Time\ 2} = 99$) to establish how rare FREs believe different minutia types to be and to determine the variation in examiners' perceptions—both between different examiners and across time for the same examiner. We observed significantly less variation in FREs' perceptions of minutia frequency for three minutiae: the two most common minutiae and the minutia perceived to be the least common. We also observed increases in FREs' estimates of minutia frequency over time and when they reported recent sightings of the rarest minutiae. FREs reported frequently using this information in their fingerprint comparison decisions. We present practical recommendations for using these consensus-based frequency estimates (until more objective data are available) to increase consistency in FREs' use of base rates when examining fingerprint evidence, which may consequently increase the repeatability and reproducibility of decisions made by FREs.

Latent fingerprint analysis is a useful forensic tool because fingerprint patterns have distinctive elements. If there are many features in agreement between two fingerprint impressions—for example, one collected from a suspect (an *exemplar* print) and one found in connection with a crime (a *latent* print)—a *latent print examiner* (LPE) can opine about the likelihood that a particular individual left a particular fingerprint impression [4]. Most forensic science disciplines have received harsh criticism over the last few decades (e.g., [15]), but there have been more validation studies on the performance of LPEs in response than seen in many other forensic sciences (e.g., [20,26,27]). However, foundational questions about fingerprint analysis remain unanswered, some of which may further shed light on the value of fingerprint comparisons and help improve the performance of, and reduce variability between, LPEs. The current research focuses on one such foundational question, addressing the very things that make fingerprint analysis so useful and compelling—the distinctive features within each fingerprint known as *minutiae*.

Friction ridge examiners (FREs; which includes LPEs and *Tenprint Examiners*/TPEs) consider the appearance, arrangement, and number of corresponding minutiae along with other information to make decisions about whether two fingerprints are likely from the same source (sometimes called an *inclusion* or, more commonly, an *identification*) or different sources (an *exclusion*). But first, FREs must opine on the quality and quantity of observed features within the fingerprints. If a fingerprint lacks sufficient quality or quantity of features to perform a comparison, an FRE may opine that the fingerprint has *no value* or is *unsuitable* for

☆ We have no known conflicts of interest to disclose. This work was funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreement 70NANB20H019 between the National Institute of Standards and Technology (NIST) and Iowa State University in the USA. Supplementary analyses and materials, R code, and preregistration information can be found at https://osf.io/xwm9h

* Correspondence to: Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada.
E-mail address: aquigley@sfu.ca (A. Quigley-McBride).
[1] https://orcid.org/0000-0001-6940-8707
[2] https://orcid.org/0000-0002-2563-438X
[3] https://orcid.org/0000-0001-5175-8960

https://doi.org/10.1016/j.forsciint.2024.112244
Received 10 March 2024; Received in revised form 26 September 2024; Accepted 6 October 2024
Available online 10 October 2024
0379-0738/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

comparison. Among fingerprints deemed suitable for comparison, the available information sometimes still does not support an identification or an exclusion decision, resulting in an *inconclusive* decision [31]. Importantly, the number of corresponding features required before an FRE will judge two impressions to be from the same source has not been standardized. FREs vary considerably in the minimum number of corresponding features they deem sufficient for a "same-source" determination, both when asked to self-report those numbers [29] and when actually performing comparisons (e.g., [17]). This variability is, in part, because there are few standardized procedures and criteria available to regulate forensic disciplines generally within the United States [15], but also because it is difficult to choose a valid cut-off value without empirical data to inform that decision.

There are two primary questions relevant to determining the number of corresponding minutiae required for a same-source determination. FREs must first determine the number of corresponding minutiae between fingerprints. In other words, knowing how many minutiae can be expected to correspond by chance in two prints that are *not* from the same source and at what point the number of corresponding minutiae suggests that two prints originate from the same source. Although a seemingly simple question, this issue has become increasingly tricky in recent years. As fingerprint databases increase in size and search algorithms improve, it is becoming more likely that FREs will observe fingerprints from different sources that share many features in common (see discussions of "*close non-matches*" in [11], and [13]). The second key question concerns the rarity of all corresponding minutiae. Logically, exceedingly common minutia types are more likely to be repeated across multiple sources due to their high prevalence. It is therefore important to determine the rarity of each corresponding minutia. If one or more of the corresponding minutiae are particularly rare, an FRE might determine that fewer corresponding minutiae are required to support an identification decision due to the small likelihood of observing such minutiae at all, let alone observing corresponding minutiae across two impressions.

1. Existing empirical data on the prevalence of minutia types

Addressing the issues raised above would require empirical knowledge of the prevalence of minutia types within fingerprints. Given the foundational importance of minutiae in latent fingerprint analysis, one might assume that these data are readily available and well-known. However, there are few examples of published, objective data that FREs can use to judge the "commonness" or "rarity" of each type of minutiae (see [12], for a review). Instead, FREs generally rely on their training and work experience to make subjective judgments about how often they have seen different types of minutiae and what the base rate of a particular minutia type is in the population.

Researchers have documented counts of basic (e.g., Bifurcations or Ridge Endings) and compound minutiae (e.g., Enclosures, Short Ridges, or Double Bifurcations) in several small-scale studies of specific populations. That said, different studies report the frequency with which different minutiae appear in different ways—some as a proportion of all minutiae counted (e.g., [10]), and some by splitting a fingerprint into manageable sections and looking at what is in those sections [19]. Combined with a lack of consistency in labelling and definitions for minutia types, it is extremely difficult to compare rates across studies.

One high-quality example of a study examining minutia frequency was conducted by Gutiérrez-Redomero and colleagues [10]. In this study, they reported counts of 18 minutia types in a sample of 200 Caucasian men and women from Spain. They found that "Ridge Endings" and "Bifurcations/Convergences" (see Fig. 1A) were the most common minutia types by far (approximately 60 % and 30 % of all minutiae found, respectively), and that complex, compound minutiae such as "Trifurcations" and "Ms" (see Fig. 1B) were among the rarest (estimated to make up approximately 1 % of all minutiae or less). Other researchers have calculated similar base rate estimates (for a review, see: [12]; for examples, see: [1][4]; [5,19]), but all that can be concluded with certainty is that some shapes are so rare that even finding one example in any fingerprint is unlikely (e.g., Trifurcations, depicted in Fig. 1B, are estimated to represent 0.05 % of minutiae by [10]), whereas other minutiae appear multiple times in almost every fingerprint (e.g., Bifurcations and Ridge Endings).

Fig. 1A. *Two Examples of Common Minutiae. Note.* These are a "Ridge Ending" and a "Bifurcation," with images obtained from Gutiérrez-Redomero et al. [10].

Fig. 1B. *Two Examples of Rare Minutiae. Note.* These are a "Trifurcation" and an "M," with images obtained from Gutiérrez-Redomero et al. [10].

2. Establishing consensus-based estimates of minutia prevalence

In the absence of conclusive data detailing the prevalence of different types of minutiae in fingerprints, researchers can nevertheless advance the forensic science discipline by examining FRE beliefs about such prevalence. Indeed, a large body of social science research indicates that the lack of empirical data does not preclude one from holding strong beliefs about a topic (e.g., [2]). As previously mentioned, FREs still hold beliefs about the prevalence of certain minutia types, but they base such beliefs on anecdotes and personal experience as opposed to empirical data.

To our knowledge, only one study has explored FRE perceptions of minutia frequency [18]. Although the findings are certainly dated (the study was conducted approximately 60 years ago), results indicated that FREs thought Ridge Endings and Bifurcations were most common and Double Bifurcations, Trifurcations, and Bridges were least common. Moreover, surveyed FREs noted that, when they relied on fewer corresponding minutiae to opine that two prints originated from the same source, it was typically because the fingerprints had "unusual characteristics." However, there was very high variability in which features examiners considered to be unusual as opposed to common. Again, these results suggest that without empirical data, FREs rely on subjective opinions to determine which minutiae are "unusual" versus "common."

3. The impact of FRE's individual estimates of minutia prevalence on analytic decisions

Assessing FRE perceptions of minutia prevalence is critical because

[4] Although we cite this study as one that reported minutiae base rates, we note that their data show some strange proportions that do not align with other data regarding base rates of minutiae. Specifically, their ridge ending counts seem very low—too low to be correct.

some minutiae carry more weight than others in the mind of the FRE. FREs assign this "weight" in a subjective manner and there is a large body of psychological literature suggesting that the weight assigned to identical minutia types is likely to vary—both between FREs and for the same FRE across time. Research shows that perceptions of relevant base rates influence FRE comparison decisions—FREs who believe identifications to be rare in a sample of prints appear more likely to "miss" same-source pairs than FREs who believe identifications to be more prevalent in the sample [9]. Findings suggest that the inverse may also be true—when FREs do not expect to encounter different-source pairs very often, they are more likely to falsely identify prints than other FREs.

Thus, perceived frequency of different minutia types might influence whether FREs conclude "identification" or "exclusion", outside of their awareness. Yet, unlike a lot of subjective information available to FREs, the frequency with which different minutiae occur is *relevant* to the task FREs perform. Issues arise not because the information is irrelevant and should not be relied upon but, rather, because of their reliance on *subjective* perceptions of frequency. If FREs' perceptions are skewed or inaccurate, this will lead to inconsistent and potentially inaccurate outcomes due to *cognitive bias*—when the collection, perception, and interpretation of decision materials, or their resulting decisions, evaluations, or confidence, is affected by an individual's preexisting beliefs, expectations, motives, or the situational context [22].

Of course, relying on perceptions of minutia prevalence is not of concern if such perceptions are accurate. Unfortunately, there is a wealth of research demonstrating that people generally struggle to accurately estimate the frequency of events or features when drawing on their own knowledge and experience [3,24,25]. People are more accurate in their estimations of base rates for events or features that are extremely rare, or very common—"*that occurs in every case*" versus "*I have almost never seen that despite my large sample of experiences*"—but everything in between is difficult to distinguish. This is because experiencing something in 30 % of relevant cases is going to feel much the same as if it occurred 40 % of the time.

People's perceptions of how often particular events or features occur are also heavily influenced by instances that are easier to bring to mind [24]. Some instances come to mind more easily because they are particularly noteworthy or memorable, such as a particularly important or controversial case, or a fingerprint with a particularly unusual set of minutiae. Instances that stand out are easier to remember and, thus, more available when people reflect on the full range and number of relevant experiences. Recent instances are also more easily available during recall [16,24]. People's beliefs about frequency are therefore unlikely to be static because perceptions rely heavily on dynamic information such as recent exposure to certain events and changing circumstances or environments. Thus, instances that occurred a long time ago are less accessible than recent instances—an FRE's experiences with minutiae last week will be easier to recall than experiences from the start of their career. Thus, memorable or recent examples are likely to be over-represented when FREs make estimates about minutiae base rates.

Other research suggests FREs also struggle to estimate base rates for different friction ridge pattern types. In one study, Mattijssen et al. [14] found that FREs' perceptions of the frequency with which different friction ridge patterns occur (e.g., Double Loop, Tented Arch, Ulnar Loop) did not align with objective counts. However, it is important to note that this study asked FREs to think about base rates for a wider range of fine-grained categories (35 different sub-categories of friction ridge patterns) than FREs typically use in their work. As a result, the accuracy of the subjective estimates in this study likely suffered compared with what the researchers would have seen if they had merely asked FREs to opine about base rates for friction ridge patterns commonly used by FREs in their work (the eight main sub-categories of friction ridge patterns).

The existing data suggest that FREs hold subjective estimates of minutia base rates, but we do not know how much variation exists in FREs' perceptions, or how well those perceptions align with actual base rates. Further, no research has addressed whether LPEs explicitly use minutia prevalence to inform their analytic decision-making. Therefore, this information may influence FREs' decisions in different ways and to different extents between FREs and across contexts. Recent work suggests that even small shifts in FREs' decision criteria can have large effects on error rates and the value of their decisions [23], and perceived minutia base rates might impact FREs' decisions at every stage of their evaluation.

The nature of the decisions and the details of the analytic procedure may vary between FREs, laboratories, and jurisdictions but, in general, the widely accepted framework for latent print examination that is followed by most FREs is *ACE-V* (*Analysis, Comparison, Evaluation,* and *Verification*; see [31], for a summary of this approach), which standardizes broad steps of the examination process. FREs' adherence to ACE-V standards is unknown [30] but research suggests that FREs' decisions vary widely within each phase of the examination process. For example, Ulery and colleagues [28] found that FREs rarely agree on how many minutiae are observed within the Analysis phase. Further, Gardner and colleagues [8] noted that FREs often differ in the types of information they review during the evaluation process.

Taken together, the existing literature suggests the ACE-V framework leaves significant room for variation and inconsistency in FRE decisions and, as shown in Table 1, one source of variation within each stage of this framework is perceived minutia base rates. The absolute or relative frequency with which particular minutiae can be expected to appear in

Table 1
How Perceptions of Minutia Rarity Affect Stages of the Procedure Used by LPEs.

Stage of Analysis		Goal	How do FREs' perceptions of how rare minutiae are factor into this stage?
A	Analysis	A decision about the value of the data in the fingermark—Is there enough reliable information in the latent fingerprint to perform a meaningful comparison with an exemplar fingerprint and support a decision about source?	Rare minutiae are more distinctive. FREs expect rare minutiae to appear less often—particularly in fingerprints from different people. Thus, FREs perceive rare minutiae to have more value, or weight, than common minutiae.
C	Comparison	An observation of the similarities and dissimilarities in the latent fingerprint and the exemplar fingerprint from a subject. Are there corresponding features? Are there dissimilarities? How clear and distinctive are the features?	If two impressions feature a minutia perceived as rare by the FRE, they may assign more weight to that similarity than to other similarities. Corresponding, perceptually similar minutiae that are rare may outweigh some dissimilarities.
E	Evaluation	The FRE makes a decision based on their analysis and comparison of the fingermark and fingerprint—Are there enough corresponding features to support an "identification", an "exclusion", or an "inconclusive" decision?	Minutiae that are perceived as rare, yet are shared by two impressions might indicate there is strong support for a "same-source" determination compared with common corresponding minutiae observed in two impressions.
V	Verification	A second FRE compares the fingermark and the fingerprint to corroborate the first FRE's decision. Ideally, this is done independently with no knowledge of the initial decision, or of the information used by the first FRE to complete their analysis.	The second FRE may have slightly different perceptions of the rarity of a particular minutia as these estimates are based on their experience, not objective data. So, the second FRE may assign a different level of value to the minutia than the first FRE.

Note. FRE = Friction Ridge Examiner.

fingerprints is also clearly task-relevant in latent print examination. That said, the relevant information can still lead to undesirable variation in FREs' decisions about the same latent print information if that information is highly suggestive of an inclusion or exclusion, inconsistently applied, or relies on subjective information (i.e., if the *biasing power* or *subjectivity* is high; [6]). Therefore, it is important to develop an understanding of variation in perceptions of minutia base rates and the extent to which this subjective information might shape FREs' decisions.

4. The current study

In the current study, we gathered information about the variability and stability of FREs' perceptions of minutia frequency. We were primarily interested in the variation in perceptions of base rates between FREs who responded to an online survey. Given the lack of empirical data to inform their opinions, we expected FREs to vary in their frequency estimates for most minutia types. Specifically, we expected a higher level of consensus among FREs' estimates of very common (e.g., Ridge Endings) and very uncommon (e.g., Trifurcations) minutia types as compared to estimates of minutiae that are neither very common nor very uncommon ([16,24]). The current study also explored the stability of FREs' perceptions, or how much a single FRE's estimates of minutia frequency change over time, by distributing a follow-up survey one month later. This second survey also examined the influence of recent experiences with rare minutiae on perceptions of base rates. We anticipated that, on average, FREs' estimates of minutia frequency would vary between time points and that FREs who reported seeing rare minutiae in the last month would report higher base rates for those minutia types than those who did not. Finally, we explored how FRs report using their perceptions of minutia base rates to inform analytic decisions. The pre-registration can be found on the Open Science Framework (https://osf.io/xwm9h/; [21]) along with other materials, data, and R Code.

5. Method

5.1. Participants

We initially recruited 132 FREs by advertising via email to a list of FREs who had previously indicated their willingness to participate in research. Ninety-nine of our initial participants also completed a second questionnaire at Time 2, one month after the first questionnaire, which resulted in a full set of data from 75 % of the sample. The average age of participants was 41.66 ($SD = 9.06$). Most FREs were female (76 %), from the USA (95 %), and regularly completing FRE casework (97 %). These FREs reported working across 29 different states in the USA, with the highest proportions of FREs in California (24 % of the sample) and Florida (8 % of the sample), followed by Virginia and Texas (6 % of the sample in each).

When asked how long they had been working as an FRE, the average response was 11.74 years ($SD = 7.64$) and, when asked how much of their work time was spent on casework, the average amount of time reported by FREs was 92 %. Most FREs reported working at an accredited forensic laboratory (78 % of the sample) and just over half were certified by the International Association for Identification (IAI; 58 % of the sample).

5.2. Design and analyses

Our data consist of repeated measures at two time points, which allows us to assess both inter- and intra-FRE variability in perceptions of minutia frequency. First, we examined rank-order judgments and frequency estimates at Time 1 to assess inter-FRE variation in perceptions of minutia frequency ($n = 132$). Then, we compared Time 1 and Time 2 responses to examine how FREs' perceptions might vary over time ($n = 99$). In our preregistration, we planned to conduct Hotelling's T analyses to compare estimates at Time 1 and Time 2. However, because we achieved a higher sample size than anticipated due to low attrition rates, we were able to run a multilevel analysis controlling for the variation attributed to the individual FRE and the variation associated with different types of minutiae at Level 2. At Level 1, we included Time and conducted several analyses to check for any lab or FRE-level factors that might also influence the stability of estimates (e.g., amount of casework). We also analyzed FREs' responses to questions about how they use their perceptions of minutia prevalence in their work and how their training and laboratory procedures have addressed the issue of minutia base rates, if at all ($n = 132$). Finally, we ran non-parametric analyses on the data regarding recent exposure to rare minutiae to account for the unequal sample sizes across the key comparison groups.

5.3. Materials

5.3.1. Minutia Images

We sought to ensure that we used representative and realistic language when referencing minutia in our survey. We did not want to cause confusion by labelling the categories of minutiae differently than what FREs usually see. Because there are no standardized labels for minutiae [7], we asked participants to estimate the prevalence of 14 basic and compound minutia types based on exemplary images of minutia as opposed to text labels alone. Table 2 contains the images used and the labels that we have assigned to them (see [7], for our reasoning about the labels). Some of the minutia images were obtained from Prof. Christophe Champod at the University of Lausanne, Switzerland, and were used with his permission. Where possible, we included two examples of minutia types so that the variation that might be seen in the shape of interest was communicated to participants. As shown in Table 2, each image depicted only one type of minutia and no others.

5.3.2. Survey questions

The first survey (hereafter, Time 1 survey) asked FREs to report basic demographic information and some general information about their work and career to date. One month later, Time 1 FREs who agreed to participate in a follow-up survey (hereafter, the Time 2 survey) were asked if anything about their work environment or the nature of their work had changed since the last survey.

The main part of the surveys asked FREs to estimate the frequency with which 14 different minutia types appear in fingerprints. FREs answered this question in two different formats (see below). They then answered questions about how they incorporated such information in their casework and identified the sources from which they had obtained their beliefs about minutiae frequency.

Ranking information. FREs ranked the images of minutiae from the most common to the least common. The main prompt was:

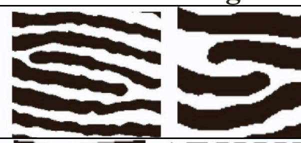
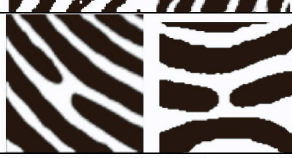
Consider the ***frequency*** *with which you come across different feature types* ***in general*** *during your work as a fingerprint/friction ridge examiner. Please rank the following minutiae by dragging the pictures around until they are in order from the* ***most common feature at the top*** *and the* ***least common feature at the bottom.***

The 14 images containing examples of each minutia type initially appeared in a random order. Participants could click and drag them to the position they felt most appropriate. They were also able to provide open-ended responses to explain any decisions they made in more detail (e.g., if two minutia types should be ranked equally common) or to make other comments.

Percentage Estimates. We next presented FREs with the same images of 14 minutiae (see Table 2) and asked that they label each minutia (see [7], for these data). Then, they were asked, *"Consider all of the fingerprints you have evaluated during your time as a fingerprint/friction ridge examiner. Approximately what percentage contained at least one example of this minutia type?"* They were asked to report their estimates again one month later, for the Time 2 survey, using the same questions.

Decision-Making Process Questions. The final phase of the survey

Table 2
Images Containing Examples of the Fourteen Minutia Types of Interest.

Minutia Images	Label	Minutia Images	Label
	Ridge Ending		Over/Under
	Bifurcation		Bridge
	Dot		Divergence
	Enclosure		Spur
	Short Ridge		Tuning Fork
	Ridge Break		Double Bifurcation
	Crossover		Trifurcation

Note. The reasoning behind our choice of labels can be found in Eldridge et al. [7], which is an article presenting data on minutia label consistency between FREs.

involved answering questions about whether the prevalence of minutia types influenced participants' analytic decisions. FREs were asked if they identified only Bifurcations and Ridge Endings, or also referred to compound or combined minutiae (e.g., Enclosures, Spurs) during their analyses. Participants were also asked to indicate what type of information they used to inform their estimates of the frequency of different minutia types. Options included published studies or other written reports, their own casework experience, their training, or conferences.

Percentage Estimates. We next presented FREs with the same images of 14 minutiae (see Table 2) and asked that they label each minutia (see [7], for these data). Then, they were asked, *"Consider all of the fingerprints you have evaluated during your time as a fingerprint/friction ridge examiner. Approximately what percentage contained at least one example of this minutia type?"* They were asked to report their estimates again one month later, for the Time 2 survey, using the same questions.

Decision-Making Process Questions. The final phase of the survey involved answering questions about whether the prevalence of minutia types influenced participants' analytic decisions. FREs were asked if they identified only Bifurcations and Ridge Endings, or also referred to compound or combined minutiae (e.g., Enclosures, Spurs) during their analyses. Participants were also asked to indicate what type of information they used to inform their estimates of the frequency of different minutia types. Options included published studies or other written reports, their own casework experience, their training, or conferences.

Then, participants were asked, "*How often does the rarity of observed minutia play a role in your examination decisions?*" followed by specific questions about value/no value decisions and comparison decisions separately.

Examples of Rare Minutia Questions. During the Time 2 survey only, FREs were asked if they had seen an example of the five rarest minutiae since the Time 1 survey—in the last month. The "rarest minutia types" were defined as the five minutia types perceived by FREs to be rarest according to base rate estimates at Time 1. The rarest minutiae, according to Time 1 participants, were Bridges, Crossovers, Divergences, Double Bifurcations, and Trifurcations (see Table 3).

5.4. Procedure

Two surveys were administered on Qualtrics. A link for the first survey was shared with potential participants via email, with two reminders sent in the second week after initial distribution. After reading the consent form and agreeing to participate, FREs provided information about themselves, the laboratory they work for, and their career. Then, they took part in a rank-order task where they ordered images of 14 different minutia types based on their perception of how common each was relative to the others.

Table 3
Summary of the Descriptive Statistics for FREs' Perceptions at Time 1.

Minutia Type	Time 1 Rank Scores		Time 1 Frequency Estimates	
	M_{Rank}	SD_{Rank}	$M_{Frequency}$	$SD_{Frequency}$
Ridge Ending	1.52	1.64	96.75 %	9.71 %
Bifurcation	2.09	1.69	96.19 %	10.95 %
Tuning Fork	4.21	2.07	64.44 %	25.52 %
Short Ridge	5.05	1.66	59.40 %	24.25 %
Dot	6.04	2.36	52.71 %	25.60 %
Over/Under	6.49	2.06	46.37 %	22.27 %
Enclosure	6.91	2.20	46.21 %	23.62 %
Divergence	8.29	2.41	43.66 %	24.14 %
Ridge Break	8.44	2.64	38.06 %	24.48 %
Spur	8.88	2.31	33.76 %	22.12 %
Crossover	10.34	2.00	23.92 %	17.35 %
Bridge	11.55	1.56	18.27 %	17.63 %
Double Bifurcation	11.67	2.66	21.26 %	20.02 %
Trifurcation	13.53	1.48	6.00 %	10.22 %

Note. M_{rank} = average rank assigned to minutiae across participants. $M_{frequency}$ = the average percentage of fingerprints in which FREs estimated they had seen this minutia type. SD_{rank} and $SD_{frequency}$ = the standard deviations for those same variables. FRE = friction ridge examiner.

Next, for each of the 14 minutia types, participants were given one or two pictures "…*of a type of fingerprint minutia to reference*" and asked "…*what name [do] you use to refer to that minutia/feature of a fingerprint*". These data are reported separately in Eldridge and colleagues article [7]. Then, participants were asked to provide a value between 0 % and 100 % to indicate what percentage of fingerprints they had reviewed contained that minutia type.

The survey finished with questions about their laboratory procedures and their approach to using their perception of minutia rarity in their fingerprint analyses and decisions. FREs were then asked if they were willing to complete a second survey containing follow-up questions about what they just reported. If they said no, they were debriefed before submitting the survey. FREs who indicated they would complete a follow-up survey were emailed approximately one month later with the link to the Time 2 survey. After reporting whether anything about their FRE work had changed since the last survey, they were again asked to provide their estimate minutia frequency. Then, they were asked to report whether they had seen an example of the five rarest minutiae since the last survey before they were debriefed.

6. Results

To answer our research questions, we examined inter-FRE variation in the ranks assigned to each minutia type and the frequency estimates at Time 1 ($n = 132$). Then, we assessed how FREs' perceptions might vary over time—intra-FRE variation ($n = 99$)—or might change after seeing an example of a rare minutia ($n = 98$). Finally, we analyzed how FREs reported using their estimates of how rare minutiae are and how their training and laboratory addresses the issue of minutiae base rates, if at all (n =132).

6.1. Inter-FRE variation

The Time 1 data included ranking information and frequency estimates for each minutia type. A summary of the descriptive statistics associated with each minutia type is presented in Table 3, and 95 % Confidence Intervals for the perceived base rates are displayed in Fig. 2.

6.1.1. Rankings

As detailed in Table 3, perceptions of rarity for several minutiae clustered at similar ranks, suggesting a lack of consensus over which is rarest among FREs or that these minutiae occur at a very similar frequency. However, there was substantial agreement in rankings within minutia type when examined using intraclass correlations (ICC = 0.79, 95 % CI [0.66, 0.92]). For example, Dots, Over/Unders, and Enclosures all received a mean ranking between 6 and 7. Divergences, Ridge Breaks, and Spurs were also grouped together with mean rankings between 8 and 9.

6.1.2. Perceived base rates

As detailed in Table 3 and visually depicted in Fig. 2, frequency estimates mirror the ranking data, with only two minutiae receiving average rankings and average perceived base rates that did not match up—Bridges (M_{rank} = 11.55; $M_{Frequency}$ = 18.27 %) and Double Bifurcations (M_{rank} = 11.67; $M_{Frequency}$ = 21.26 %). The Time 1 frequency estimates (summarized in Table 3) reveal something that the ranking data did not, though. There are two minutiae that are seen in almost every fingerprint: Ridge Endings ($M_{Frequency}$ = 96.75 %) and Bifurcations ($M_{Frequency}$ = 96.19 %). However, the minutia with the third highest perceived base rate was a Tuning Fork, for which the average perceived base rate was 64.44 %. The average frequency estimates appear to drop at a fairly even pace until we reach the least common minutia type—Trifurcations ($M_{Frequency}$ = 6.00 %), which is a 12 % drop from the average for Bridges ($M_{Frequency}$ = 18.27 %). When examining the intraclass correlation for the frequency estimates, agreement among FREs was lower than observed in the ranking data, but still moderate (ICC = 0.60, 95 % CI [0.42, 0.79]).

There were also patterns observed in the variation of these estimates. The minutiae perceived to be the most common (Ridge Endings and

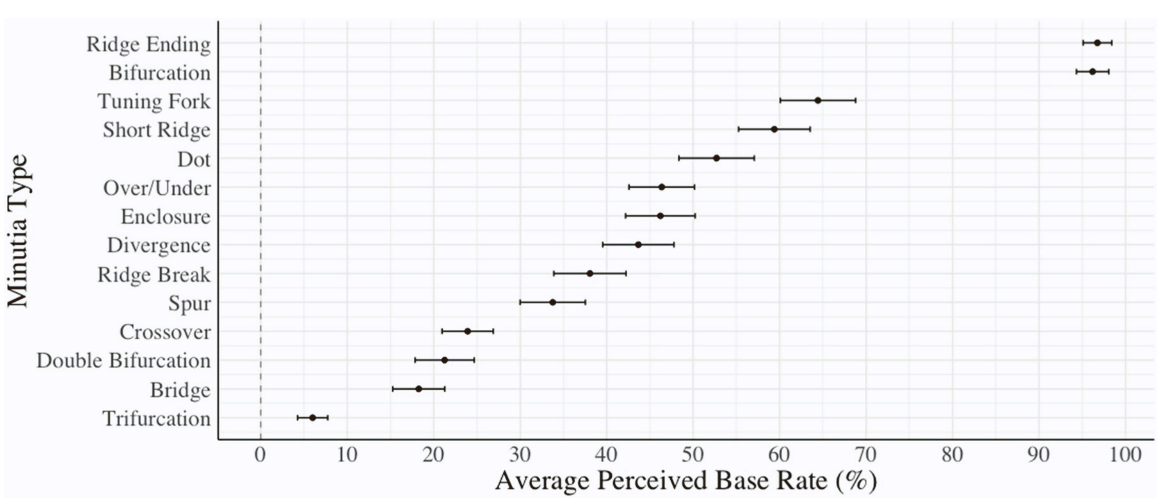


Fig. 2. *Average Perceived Frequency of 14 Minutiae Among Friction Ridge Examiners* (N = 132; Error bars are 95 % Confidence Intervals).

Bifurcations) and the rarest (Trifurcations) were associated with less variation in estimates than seen for other minutia types ($SD_{frequency}$ range from 9.71 % to 10.22 %). Whether there were significantly different standard deviations associated with different minutia types was analyzed by identifying Ridge Ending estimates as the reference group (the minutia with the smallest amount of variation). Other minutia frequency estimates were compared to this reference group using ratios.

If two standard deviations are very similar, the resulting ratio will be close to 1 and the 95 % Confidence Interval will include 1. If they are significantly different, the 95 % CI will be completely above or below 1. We looked for standard deviations that were significantly higher than the standard deviation associated with FREs' Ridge Ending frequency estimates—where the 95 % Confidence Interval for the ratio was entirely below 1 and the *p* value for the *F*-test was below 0.05. Results are displayed in Table 4 and indicate that there was significantly less variability FREs' perceptions of Ridge Endings, Bifurcations, and Trifurcations than for all other minutia types. It is notable that these minutiae are the three most/least common minutia types as perceived by FREs (the most extreme perceived base rates). The variation between perceived base rates is shown in Fig. 2 and the results of the inferential analyses are reported in Table 4.

6.2. Intra-FRE variation

A total of 99 FREs completed the frequency estimates again at Time 2 (75 % of the original sample), and 98 of these FREs completed questions about whether they had seen the five rarest minutiae in the time since the first survey. The descriptive statistics for Time 2 base rate estimates from 99 FREs who completed both surveys are displayed in Table 5. Fig. 3 shows the extent to which the mean frequency estimates for each minutia type at Time 1 and Time 2 were consistent or inconsistent.

A Multilevel Multiple Regression model was run with FRE and Minutia Type as Level 2 effects so that the variance explained by individual tendencies and by the different minutia types was controlled for in the model. The first model included lab accreditation status and casework status as main effects in addition to assessing the main effect of time on estimates of minutia frequency, which revealed a significant effect of time (B =.06, p <.001) and FRE casework status (B =.10, p =.007), but the lab accreditation status variable did not reach significance ($B = -0.07$, p =.058). Finally, a model was run assessing potential interaction effects among these three variables. There was no three-way interaction detected between these variables ($B = -.15$, p =.393), so a reduced model was run with two two-way interactions (Time*Lab

Table 4
Comparing the Standard Deviations of FREs' Estimates Between Minutia Types at Time 1.

Minutia Type	$SD_{Frequency}$	F value	p value	95 % CI – Ratio of Variances*
Ridge Ending	*9.71 %*	*Reference Group*		
Bifurcation	10.95 %	0.79	0.17	[0.56, 1.11]
Tuning Fork	25.52 %	0.14	<.001	[0.10, 0.20]
Short Ridge	24.25 %	0.16	<.001	[0.11, 0.23]
Dot	25.60 %	0.14	<.001	[0.10, 0.20]
Over/Under	22.27 %	0.19	<.001	[0.13, 0.27]
Enclosure	23.62 %	0.17	<.001	[0.12, 0.24]
Divergence	24.14 %	0.16	<.001	[0.11, 0.23]
Ridge Break	24.48 %	0.16	<.001	[0.11, 0.22]
Spur	22.12 %	0.19	<.001	[0.14, 0.27]
Crossover	17.35 %	0.31	<.001	[0.22, 0.44]
Bridge	17.63 %	0.30	<.001	[0.22, 0.43]
Double Bifurcation	20.02 %	0.24	<.001	[0.17, 0.33]
Trifurcation	10.22 %	0.90	0.557	[0.64, 1.27]

Note. $N = 132$. * The *SDs* are compared using a ratio. A SD that is significantly higher than the standard deviation associated with estimates of Ridge Ending prevalence (reference group and the smallest SD), the 95 % Confidence Interval will be below 1 and the *p* value below 0.05.

Table 5
Summary of Descriptive Statistics for FRE Perceptions at Time 2.

Minutia Type	Time 2 Frequency Estimates	
	$M_{Frequency}$	$SD_{Frequency}$
Ridge Ending	97.38 %	7.51 %
Bifurcation	96.47 %	9.40 %
Tuning Fork	73.32 %	24.83 %
Short Ridge	66.02 %	25.46 %
Dot	59.77 %	25.74 %
Over/Under	52.59 %*	25.64 %
Enclosure	52.72 %*	24.78 %
Divergence	48.08 %	25.50 %
Ridge Break	44.48 %	27.28 %
Spur	42.95 %	23.26 %
Crossover	28.60 %	23.04 %
Bridge	24.38 %*	20.23 %
Double Bifurcation	21.17 %*	23.23 %
Trifurcation	8.92 %	14.64 %

Note. $N = 99$; FRE = Friction Ridge Examiner.

Accreditation Status and Time*FRE Casework Status) instead. There was a significant interaction effect between lab accreditation status and time on minutia frequency estimates ($B = -0.12$, p =.001) but no significant interaction between FRE casework status and time (B =.10, p =.103). See Table 6 for a summary of these results.

Recent exposure to rare minutiae. We asked participants if they had encountered any of the five minutiae judged to be the rarest, on average, at Time 1. Table 8 breaks down the participants' estimated base rates at Time 2 based on whether they reported seeing an example of these minutiae in the time since the first survey. Some are very rare, so few FREs reported seeing an example of one since the first survey (e.g., only 11 people reported seeing a Trifurcation). However, others were more common, so few people reported that they did *not* recall seeing an example since the first survey (e.g., only 8 people reported that they did *not* encounter a Divergence since the Time 1 survey).

Due to the uneven sample sizes in the groups relevant to this comparison, nonparametric inferential statistics (a Kruskal-Wallis nonparametric ANOVA) were calculated to determine whether FREs who observed one of five rare minutiae reported higher prevalence rates. There are several effects to note here that are summarized by the aggregate descriptive statistics in Table 7 and the breakdown by minutia type in Table 8. There was a significant difference in estimates between examiners who reported seeing rare minutiae and those who did not, even at Time 1 (M = 30.91 %), with those who reported seeing rare minutiae estimating higher frequencies across the board. However, we also observed even higher estimates of base rates for rare minutiae at Time 2 (M = 36.10 %) among FREs reported seeing those rare minutiae recently, whereas there was no significant different between Time 1 (M = 12.03 %) and 2 (M = 13.70 %) for those who did not; $\chi^2(N = 98, 3) = 235.1$, p<.001. Thus, as expected, we observed an effect associated with recent exposure to rare minutiae.

How often did FREs report using their perceptions of minutia frequency in their fingerprint examination decisions? At Time 1, FREs were asked about their use of minutia rarity in decisions about fingerprints. Most FREs reported that they overtly used their perceptions of minutia rarity when making "value/no value" decisions (84 %) and when performing comparisons (89 %). When asked how much they considered minutia frequency in general when evaluating fingerprints, most reported that they considered this information "often" (36 %) or "some" (34 %), with slightly fewer FREs selecting "always" (20 %), and very few FREs reporting "rarely" (8 %) or "not at all" (2 %).

7. Discussion

The current study provides consensus-based estimates of the base rates FREs associate with different minutia types, as well as information about when and to what extent FREs' perceptions of minutia base rates

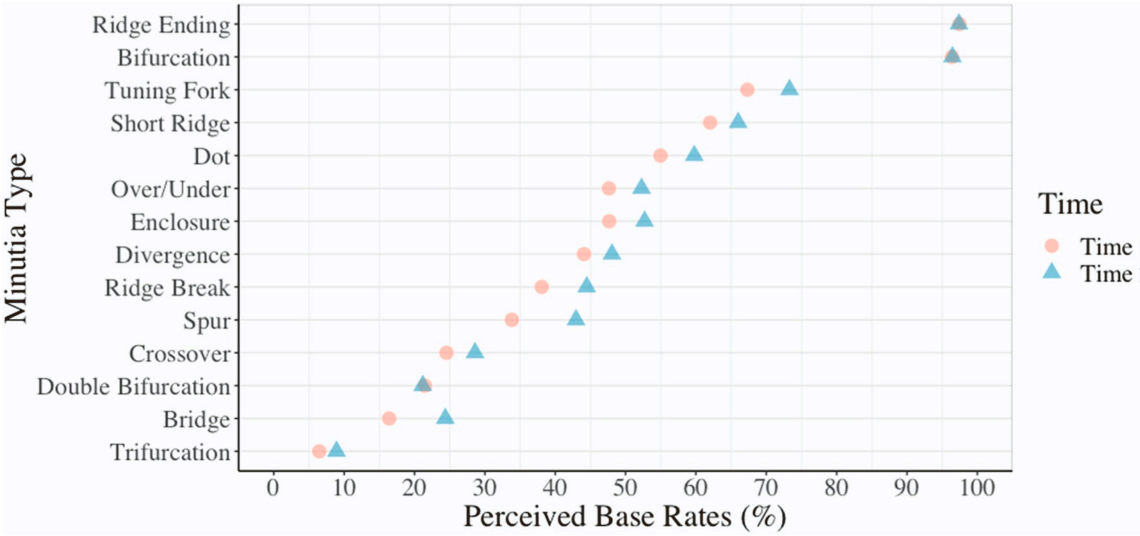


Fig. 3. *Difference in Perceived Frequency of Minutiae Among Friction Ridge Examiners at Time 1 and 2* (N = 99).

Table 6
Summary of the Main Effects and Interaction Multilevel Multiple Regression Models Assessing the Effect of Time and FRE Characteristics on FREs Minutia Frequency Estimates.

Model	Parameter	Standardized Beta	*p* value	Cohen's *d*
Main Effects	Time	0.06	<.001	0.24
	Lab Accreditation Status	−0.07	0.058	−0.39
	LPE Casework Status	0.10	0.007	0.56
Interaction	Time	0.04	0.415	0.04
	Lab Accreditation Status	0.03	0.570	0.07
	LPE Casework Status	0.05	0.289	0.13
	Time*Lab Accreditation Status	−0.12	0.001	−0.12
	Time*LPE Casework Status	0.10	0.103	0.06

Note. Participant/Minutia Type controlled for – included at Level 2 in the Multilevel Multiple Regression Analysis. Null/Intercept Model significant at p<.001. FRE = Friction Ridge Examiner.

Table 7
Descriptive Statistics Associated with the Effect of Recent Exposure to Rare Minutiae and the Timing of the Survey on FREs' Estimates of Rare Minutia Frequency.

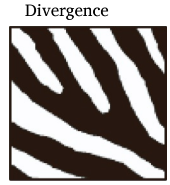
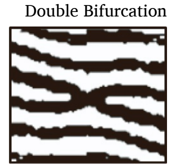
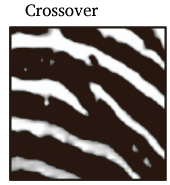
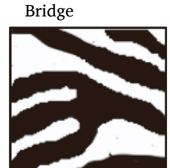
Saw Rare Minutiae?	Time	*M*	*SD*
Yes	1	30.91 %	23.89 %
	2	36.10 %	25.34 %
No	1	12.03 %	15.90 %
	2	13.70 %	18.17 %

Note. N = 98. FRE = Friction Ridge Examiner. Percentage values represent the average estimated frequency reported by FREs aggregated across the five rarest minutia types to examine the overall effect of exposure to rare minutiae between Time 1 and 2, factoring in FREs baseline estimates from Time 1.

vary. We observed some consensus between FREs regarding their perceptions of the rarity of minutiae, particularly when assessing minutia types that were extremely rare (e.g., Trifurcations) or common (e.g., Ridge Endings). Moreover, FREs exhibited more consistency in their estimates of relative minutia frequency (rankings) compared to absolute estimates (frequencies).

That said, as anticipated, we found notable variation in the perceived frequency reported by FREs for all minutia types, particularly in their estimates of minutia base rates for those friction ridge features that were

Table 8
Summary of the Estimated Frequencies Associated with FREs who Did and Did Not Encounter a Rare Minutia Between Time 1 and Time 2.

Rare Minutia Type	**Did the FRE report that they encountered an example of this minutia type since Time 1? (*n*)**	**What was the average estimated base rate for this feature among these FREs? (*SD*)**
 Divergence	Yes (*n* = 90)	49.41 % (25.88 %)
	No (*n* = 8)	33.13 % (14.62 %)
 Double Bifurcation	Yes (*n* = 39)	30.23 % (23.99 %)
	No (*n* = 59)	15.18 % (20.82 %)
 Crossover	Yes (*n* = 70)	32.89 % (23.17 %)
	No (*n* = 28)	17.87 % (19.16 %)
 Bridge	Yes (*n* = 64)	27.15 % (24.31 %)
	No (*n* = 34)	19.15 % (19.98 %)
 Trifurcation	Yes (*n* = 11)	20.55 % (24.31 %)
	No (*n* = 87)	7.45 % (12.39 %)

Note. 98 is the total number of FREs here as one participant did not complete these questions. FRE = Friction Ridge Examiner.

perceived as moderately common (e.g., Enclosures, Short Ridges). In other words, there was significantly less variation observed when FREs estimated how often they encountered the most common minutia types (Ridge Endings and Bifurcations) and the least common minutia type (Trifurcations). This is consistent with what is understood about how people estimate base rates in the absence of objective information—they draw on their personal experience, training, and practice, and what examples do and do not come to mind easily [3,24,25]. Thus, because each FRE based their estimate on different experiences, training, and pre-existing tendencies, their perceptions of the patterns in their casework also differed.

In addition to variation *between* FREs, we also found evidence of changes in FREs' perceptions of minutia frequency over time. In other words, when examining data from the *same* FRE one month later, we found that the estimated base rates reported by FREs at Time 2 were somewhat consistent with those from Time 1, particularly for those minutiae at the extremes. That is, Ridge Endings, Bifurcations, Double Bifurcations, and Trifurcations all received estimates of frequency that were almost identical across surveys (see overlapping points in Fig. 3). This suggests that FREs as a community have fairly stable perceptions of how rare these minutiae are, further supported by evidence that these minutiae were associated with the lowest standard deviations (between FREs) at Time 1. The other minutia types, though, differed in a consistent way—FREs thought that these minutia types were more common the second time they estimated the frequency with which they had seen them in fingerprints over the course of their career, resulting in a significant increase in the average estimated minutiae base rates at Time 2 (a difference of approximately 5 %; see Table 6).

The increase in estimates at Time 2 was not driven by FREs with less experience or less time spent on casework—this effect was observed even when controlling for the variation attributable to each FRE or the minutia type. That said, we suspect that some of this variation was driven, at least in part, by FREs' participation in our study. By participating in the Time 1 study, the concept of minutia frequency would have been made more salient to FREs, so they might have attended to the frequency with which different minutiae occurred more closely between the Time 1 and Time 2 surveys. Rather than discounting our results, though, this explanation provides further evidence to support the idea that FREs' perceptions of base rates are malleable.

In addition, when FREs reported encountering one of the five rarest minutiae since the first survey (Time 1), they reported even higher base rates for those minutia types on average, over and above the overall increase observed in estimates provided at Time 2—sometimes substantially higher base rates. Although changes in perceptions tended to be quite small when comparing Time 1 and Time 2 overall, the differences were much larger when focused on the rare minutiae that FREs reported seeing or not seeing since Time 1. Thus, even though the FREs were making decisions about features and events that they had been observing for years, their most recent experiences weighed heavily in their estimates about minutia frequency. We note, though, that such effects are typical when people are asked to provide subjective base rates [24], so these results suggest that FREs' perceptions of base rates are not especially unique and may be affected by processes shown to affect most heuristic decision-making processes.

Finally, when we asked FREs how they use their subjective sense of minutia rarity, most examiners (i.e., 90 %) indicated that they used such information at least sometimes in their work, with 56 % indicating that they used such information "often" or "always". These data shed light on potential sources of variation in FREs' decisions and show that examiners used these variable perceived base rates to inform their Analyses and Comparisons. Given that the appearance, arrangement, and number of minutiae inform latent fingerprint examinations, fingerprint evidence is currently and regularly affected by FREs' varied and malleable estimates of minutia frequency (refer to Table 1 for a summary of how minutia rarity impacts each stage of a typical latent fingerprint examination). Thus, perceptions of minutiae frequency and the extent to which there is consensus among FREs about the correct base rate to use will matter, and establishing a set of data that can be used to standardize the weight assigned to different minutiae would be valuable.

It is important, however, that these findings are not interpreted as a critique of FREs' skills or competence. In fact, one could argue that these data show that FREs' perceptions are remarkably consistent in a lot of ways, especially considering the lack of empirical data on which to base opinions. Nevertheless, examiners are subject to normal cognitive limitations and rely on cognitive shortcuts just like everyone else. Forming a subjective sense of how often relevant events or features occur and using that to inform decisions is a completely normal decision-making process that people engage in automatically [16]. Problems arise when there are no empirical data to understand the nature of these cognitive heuristics in a particular context, and no laboratory procedures to limit discrepant practices across examiners. Though subjective base rates are very helpful for making efficient decisions in most everyday contexts, they should not form the basis of decisions that are advertised as objective or where the consequences of error are serious [25].

These data represent the first comprehensive analysis of the variation in FREs' perceptions of minutia frequency and suggest that something should be done to reduce the variability in these perceptions among FREs. One way to prevent FREs from relying on inconsistent, subjective estimates of minutiae frequency in their work would be to ask that FREs ignore the rarity of minutiae in their analyses and comparisons. For example, many laboratories in Europe provide a fixed threshold for examiners to use when examining fingerprints—a fixed number of minutiae of any type that must be observed prior to determining that a latent fingerprint is suitable for analysis or from the same source as another fingerprint impression. This reduces the variability in decisions, but fails to capitalize on observable, valuable information. Accurate knowledge of minutia base rates could help FREs be more precise about the probative value associated with the evidence they are examining and make decisions about fingerprint comparisons that are supported by quantitative data. Thus, we do not recommend that FREs or their respective laboratories move to a "quantity-only" approach to examining minutiae because empirical research documenting base rates of minutiae has the potential to increase both the accuracy, reliability, and utility of FRE decisions.

Although empirically-supported base rate estimates for every minutia type are not yet available, there are two ways that the current findings can help improve reproducibility and repeatability in FRE decisions by reducing variation in how examiners account for their perceptions of minutia base rates. First, consensus-based lists of estimated minutia base rates could be used by FREs during training, and when conducting analyses and comparisons. Although such lists would not represent *objective* base rates derived from an actual count of minutiae from a large number of fingerprints, current findings employ the knowledge of more than 100 experienced FREs. If FREs used these consensus-based values when examining fingerprints rather than their subjective sense of minutia rarity, the use of minutia base rates in casework would be more *consistent* even if they are not totally aligned with the *actual* base rates of these minutiae. These average estimates could be incorporated into FRE training, textbooks, software, or continuing education efforts.

We also note, though, that the absolute estimates of minutia frequency varied more so than relative estimates (i.e., rank data) and therefore might not be the most effective to use in practice. In addition, the Time 2 data suggested that merely alerting FREs to the concept of minutia rarity might have inflated their base rate estimates. Another way to use the current data would be to use a "relative frequency" approach. In Table 9, we provide simple categories to guide the use of the consensus information in casework. Minutia types are categorized as very common (estimated to appear in 90 % to 100 % of fingerprints), somewhat common (50 % to 70 %), somewhat rare (30 to 50 %), rare (10 % to 30 %), and very rare (estimated to appear in less than 10 % of fingerprints) based on the ranking and frequency data obtained in the

Table 9
Minutia Types Categorized by Perceived Frequency and Rank Order.

General Perceived Frequency	Approx. Estimated Frequency Range	Minutia Types
Very common	90 % - 100 %	Ridge Ending, Bifurcation
Common	70 % - 90 %	None
Somewhat common	50 % - 70 %	Tuning Fork, Short Ridge, Dot
Somewhat rare	30 % - 50 %	Over/Under, Enclosure, Divergence, Break, Spur
Rare	10 % - 30 %	Crossover, Double Bifurcation, Bridge
Very rare	0 % - 10 %	Trifurcation

current study. We note that there are no minutiae categorized as "common" as there were no minutiae with average frequency estimates in the 70 to 90 % range.

Again, using the categories in Table 9 would improve consistency between FREs even though these categories are not based on actual base rates obtained from counts of minutia types. Because the consensus-based estimates are more varied than the data based on relative judgments, the use of ranked categories like this might be easier to implement than a quantitative base rate value, especially given that the estimated values may end up being very different from the true base rates for each minutia type and, thus, imply a level of precision that is not accurate. Regardless of whether the consensus-based frequency estimates or these evidence-based minutia frequency categories are used, these data will help encourage repeatable, reproducible, data-driven decisions about latent print comparisons by standardizing FREs' approach to using minutia frequency in their decisions.

Of course, as already mentioned, it would be better if FREs could be given access to *actual* base rates for common minutia types specific to their jurisdiction of practice. FREs reported using their perceptions of minutia rarity often, but there was variation observed in their estimates—both within and between FRE. This highlights the importance of efforts to determine the actual base rates associated with different minutia types. Consider a situation where a particular minutia type is perceived to be very rare. That feature will be assigned more weight by FREs than other minutia believed to be more common and, as a result, FREs might change their "suitable for analysis" criteria (the *Analysis* phases of ACE-V) or their criteria concerning the number of corresponding minutiae that should be present before making a same-source decision (*Comparison* and *Evaluation* phases of ACE-V). In other words, the perception that an observed minutia is rare might result in FREs requiring fewer useable minutiae to decide a latent print is "suitable for analysis" or fewer corresponding minutiae if one of those corresponding minutiae is believed to be rare. We know that the number of minutiae required for a same-source decision is subjective and varies between FREs [29]. Introducing actual base rates for common minutia types into FRE training and casework would help standardize the weight assigned to each pair of minutiae that correspond and reduce the variability in decision criteria [23]. Hence, currently underway is the next step in our planned research program—developing easy-to-use standardized protocols for determining actual base rates and obtaining counts of minutia types from a large sample of friction ridge impressions.

The current research represents a step forward in our understanding of FRE decision-making and judgments. With these data, we can better understand the sources of variation in FREs' decisions and attempt to minimize unwanted sources of variation. As is the case for any one piece of research, though, there are several limitations for the current work. We only asked FREs to estimate the frequency with which they encounter 14 minutia types because we only wanted to ask FREs to opine about minutiae that they were likely to have considered before and already viewed as a distinct sub-category of friction ridge events. So, although our data might not shed light on all possible compound minutiae, we can be more certain that our data represent experience-based estimates, not guesses. Second, we did not want the study to take too long or feel too arduous. It was important to us that we maintain the interest and support of our FRE participants so that they would be willing to complete the Time 2 survey and future studies. Another limitation is our reliance on self-reporting—we did not validate that FREs who reported seeing rare minutiae actually saw rare minutiae in the month between surveys. So, we do not know if any differences associated with those self-reports can be attributed to FREs' recent experiences with rare minutiae.

Another consideration is that our sample size was fairly small compared to the total number of FREs currently working. That said, our sample appears to be representative of the wide range of laboratories, training backgrounds, jurisdictions, and experience levels seen in North America and our numbers are in line with the numbers of FRE participants that have been garnered in other recent latent print studies [17, 26]. In addition, the demographic makeup of our sample tends to align with what is seen when surveying FREs who are associated with the International Association of Identification (IAI) suggesting our sample is generally representative of the FRE population in North America (e.g., [8]). There was also some attrition associated with collecting data at two time points. This was minimized as much as possible, but there was still 25 % attrition observed at Time 2. That said, the groups of participants at each time point did not differ in any significant way.

This study provides consensus-based data regarding the prevalence of particular friction ridge events and the extent to which FREs' perceptions of minutiae base rates vary between FREs and over time. We show that, although there is some agreement about the prevalence of minutia types—particularly when those minutiae are very rare or very common, or when examining rank data—different FREs examining the same friction ridge impression might assign different weights to minutiae observed during the Analysis and Comparison phases of fingerprint examination. FREs' judgments about minutia frequency are also dynamic—we observed changes in their perceptions of minutia base rates after only a short period of time. These data can be used to increase consistency in fingerprint examination within and between FREs by providing examiners with a consensus-based information to use when evaluating the value of a latent fingerprint or the similarity between two fingerprints rather than leaving them to rely on their own subjective perception and experience.

Funding Source Declaration

This work was funded (or partially funded) by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreements 70NANB15H176 and 70NANB20H019 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College and University of Nebraska, Lincoln.

CRediT authorship contribution statement

Brett O. Gardner: Writing – review & editing, Methodology, Conceptualization. **Heidi Eldridge:** Writing – review & editing, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Adele Quigley-McBride:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of Competing Interest

None to report.

References

[1] U.U. Akpan, T.E. Awe, D. Idowu, Types and frequency of fingerprint minutiae in individuals of Igbo and Yoruba ethnic groups of Nigeria, Ruhuna Journal of Science *10* (1) (2019).

[2] C.A. Anderson, Abstract and Concrete Data in the Conservatism of Social Theories: When Weak Data Lead to Unshakeable Beliefs, Journal of Experimental Social Psychology *19* (2) (1983) 93–108, https://doi.org/10.1016%2F0022-1031%2883%2990031-8.

[3] M. Bar-Hillel, The base-rate fallacy in probability judgments, Acta Psychologica 44 (3) (1980) 211–233, https://doi.org/10.1016/0001-6918(80)90046-3.

[4] C. Champod, A. Biedermann, Overview and Meaning of Identification/Individualization, in: M.M. Houck (Ed.), In *Encyclopedia of Forensic Sciences*, Elsevier, Oxford, 2023, pp. 53–62.

[5] C. Champod, C. Lennard, P. Margot, M. Stoilovic, Fingerprints and other ridge skin impressions, CRC Press, Boca Raton, FL., USA, 2004.

[6] I.E. Dror, J. Kukucka, 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* (2021) 100161, https://doi.org/10.1016/j.fsisyn.2021.100161.

[7] H. Eldridge, A. Quigley-McBride, B.O. Gardner, A survey of naming conventions for different minutia types in friction ridge examination, Journal of Forensic Identification *74* (3) (2024) 251–281.

[8] B.O. Gardner, M. Neuman, S. Kelley, What types of information can and do latent print examiners review? A survey of practicing examiners, Forensic Science International *344* (2023) 111598, https://doi.org/10.1016/j.forsciint.2023.111598.

[9] B. Growns, J. Kukucka, The prevalence effect in fingerprint identification: Match and non-match base-rates impact misses and false alarms, Applied Cognitive Psychology *35* (3) (2021) 751–760, https://doi.org/10.1002/acp.3800.

[10] E. Gutiérrez-Redomero, C. Alonso-Rodríguez, L.E. Hernández-Hurtado, J. L. Rodríguez-Villalba, Distribution of the minutiae in the fingerprints of a sample of the Spanish population, Forensic Science International *208* (1-3) (2011) 79–90, https://doi.org/10.1016/j.forsciint.2010.11.006.

[11] J.J. Koehler, S. Liu, Fingerprint error rate on close non-matches, Journal of Forensic Sciences *66* (1) (2021) 129–134, https://doi.org/10.1111/1556-4029.14580.

[12] G. Langenburg, Scientific research supporting the foundations of friction ridge examinations, in: E.H. Holder Jr., L.O. Robinson, J.H. Laub (Eds.), *The Fingerprint Sourcebook* (Chapter 14). U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington DC, USA, 2011.

[13] K. Li, D. Wu, L. Ai, Y. Luo, The influence of Close Non-Match fingerprints similar in delta regions of whorls on fingerprint identification, Journal of Forensic Sciences *66* (4) (2021) 1482–1494, https://doi.org/10.1111/1556-4029.14698.

[14] E.J. Mattijssen, C.L. Witteman, C.E. Berger, R.D. Stoel, Assessing the frequency of general fingerprint patterns by fingerprint examiners and novices, Forensic Science International *313* (2020) 110347, https://doi.org/10.1016/j.forsciint.2020.110347.

[15] National Research Council, Strengthening forensic science in the United States: a path forward, National Academies Press, Washington D.C., USA, 2009.

[16] H. Neth, G. Gigerenzer, Heuristics: Tools for an uncertain world, in: In.R. Scott, S. Kosslyn (Eds.), Emerging Trends in the Social and Behavioral Sciences, Wiley Online Library, 2015, pp. 1–18.

[17] C. Neumann, C. Champod, M. Yoo, T. Genessay, G. Langenburg, Improving the understanding and the reliability of the concept of sufficiency in friction ridge examination, U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington D.C., USA, 2013.

[18] J.W. Osterburg, An inquiry into the nature of proof: The identity of fingerprints, Journal of Forensic Sciences *9* (4) (1964) 413–427.

[19] J.W. Osterburg, T. Parthasarathy, T.E.S. Raghavan, S.L. Sclove, Development of a mathematical formula for the calculation of fingerprint probabilities based on individual characteristics, Journal of the American Statistical Association *72* (360a) (1977) 772–778, https://doi.org/10.1080/01621459.1977.10479954.

[20] President's Council of Advisors on Science and Technology (2016, September). Forensic science in criminal courts: Ensuring scientific validity of feature-comparison methods. Executive Office of the President of the United States. Retrieved from: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.

[21] A. Quigley-McBride, H. Eldridge, B.O. Gardner, Examiner consistency in perceptions of fingerprint minutia rarity, October 10, Open Science Framework (2024), https://doi.org/10.17605/OSF.IO/XWM9H.

[22] B.A. Spellman, H. Eldridge, P. Bieber, Challenges to reasoning in forensic science decisions, Forensic science international: Synergy *4* (2022) 100200, https://doi.org/10.1016/j.fsisyn.2021.100200.

[23] W.C. Thompson, Shifting decision thresholds can undermine the probative value and legal utility of forensic pattern-matching evidence, e2301844120, Proceedings of the National Academy of Sciences *120* (41) (2023), https://doi.org/10.1073/pnas.2301844120.

[24] A. Tversky, D. Kahneman, Availability: A heuristic for judging frequency and probability, *Cognitive Psychology 5* (2) (1973) 207–232, https://doi.org/10.1016%2F0010-0285%2873%2990033-9.

[25] A. Tversky, D. Kahneman, Judgment under uncertainty: Heuristics and biases, Science *185* (4157) (1974) 1124–1131, https://doi.org/10.1126/science.185.4157.1124.

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

[27] B.T. Ulery, R.A. Hicklin, J. Buscaglia, M.A. Roberts, Repeatability and reproducibility of decisions by latent fingerprint examiners, PloS one *7* (3) (2012) e32800, https://doi.org/10.1371/journal.pone.0032800.

[28] B.T. Ulery, R.A. Hicklin, G.I. Kiebuzinski, M.A. Roberts, J. Buscaglia, Understanding the sufficiency of information for latent fingerprint value determinations, Forensic Science International *230* (2013) 99–106, https://doi.org/10.1016/j.forsciint.2013.01.012.

[29] B.T. Ulery, R.A. Hicklin, M.A. Roberts, J. Buscaglia, Measuring what latent fingerprint examiners consider sufficient information for individualization determinations, PloS one *9* (11) (2014) e110179, https://doi.org/10.1371/journal.pone.0110179.

[30] B.T. Ulery, R.A. Hicklin, M.A. Roberts, J. Buscaglia, Changes in latent fingerprint examiners' markup between analysis and comparison, Forensic Science International *247* (2015) 54–61, https://doi.org/10.1016/j.forsciint.2014.11.021.

[31] J.R. Vanderkolk, Examination process, in: E.H. Holder Jr., L.O. Robinson, J.H. Laub (Eds.), *The Fingerprint Sourcebook* (Chapter 9), U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington DC, USA, 2011.