4-7-2021

Mt. Everest—we are going to lose many: a survey of fingerprint examiners' attitudes towards probabilistic reporting

H. Swofford
University of Lausanne

S. Cole
University of California, Irvine

V. King
University of California, Irvine

Mt. Everest—we are going to lose many: a survey of fingerprint examiners' attitudes towards probabilistic reporting

Abstract

Over the past decade, with increasing scientific scrutiny on forensic reporting practices, there have been several efforts to introduce statistical thinking and probabilistic reasoning into forensic practice. These efforts have been met with mixed reactions—a common one being scepticism, or downright hostility, towards this objective. For probabilistic reasoning to be adopted in forensic practice, more than statistical knowledge will be necessary. Social scientific knowledge will be critical to effectively understand the sources of concern and barriers to implementation. This study reports the findings of a survey of forensic fingerprint examiners about reporting practices across the discipline and practitioners' attitudes and characterizations of probabilistic reporting. Overall, despite its adoption by a small number of practitioners, community-wide adoption of probabilistic reporting in the friction ridge discipline faces challenges. We found that almost no respondents currently report probabilistically. Perhaps more surprisingly, most respondents who claimed to report probabilistically, in fact, do not. Furthermore, we found that two-thirds of respondents perceive probabilistic reporting as 'inappropriate'—their most common concern being that defence attorneys would take advantage of uncertainty or that probabilistic reports would mislead, or be misunderstood by, other criminal justice system actors. If probabilistic reporting is to be adopted, much work is still needed to better educate practitioners on the importance and utility of probabilistic reasoning in order to facilitate a path towards improved reporting practices.

Disciplines

Legal Studies

Comments

The following is a manuscript of an article published as Swofford, H., S. Cole, and V. King. "Mt. Everest—we are going to lose many: a survey of fingerprint examiners' attitudes towards probabilistic reporting." *Law, Probability And Risk* (2021). Posted with permission of CSAFE.

doi:10.1093/lpr/mgab003

Mt. Everest—we are going to lose many: a survey of fingerprint examiners' attitudes towards probabilistic reporting

H. Swofford
School of Criminal Justice, Forensic Science Institute, University of Lausanne, Switzerland

S. Cole† and V. King
Department of Criminology, Law & Society, University of California, Irvine, CA, USA

Over the past decade, with increasing scientific scrutiny on forensic reporting practices, there have been several efforts to introduce statistical thinking and probabilistic reasoning into forensic practice. These efforts have been met with mixed reactions—a common one being scepticism, or downright hostility, towards this objective. For probabilistic reasoning to be adopted in forensic practice, more than statistical knowledge will be necessary. Social scientific knowledge will be critical to effectively understand the sources of concern and barriers to implementation. This study reports the findings of a survey of forensic fingerprint examiners about reporting practices across the discipline and practitioners' attitudes and characterizations of probabilistic reporting. Overall, despite its adoption by a small number of practitioners, community-wide adoption of probabilistic reporting in the friction ridge discipline faces challenges. We found that almost no respondents currently report probabilistically. Perhaps more surprisingly, most respondents who claimed to report probabilistically, in fact, do not. Furthermore, we found that two-thirds of respondents perceive probabilistic reporting as 'inappropriate'—their most common concern being that defence attorneys would take advantage of uncertainty or that probabilistic reports would mislead, or be misunderstood by, other criminal justice system actors. If probabilistic reporting is to be adopted, much work is still needed to better educate practitioners on the importance and utility of probabilistic reasoning in order to facilitate a path towards improved reporting practices.

Keywords: reporting; testimony; fingerprint; categoric; probability; attitudes

1. Introduction

Recent years have witnessed increasing efforts to introduce probabilistic reasoning into forensic practice—particularly in the pattern evidence disciplines. We define probabilistic reasoning in forensic practice as formally recognizing and articulating the uncertainties inherent in forensic interpretation using probabilistic logic. Forensic statisticians' efforts in these areas have primarily, and understandably, been concentrated in their area of technical expertise: statistics. Therefore, these

†Email: scole@uci.edu

efforts have been focused on such activities as developing statistical models (Neumann *et al.*, 2006, 2007, 2012, 2015; Egli *et al.*, 2007, 2014; Leegwater *et al.*, 2017; Swofford *et al.*, 2018), building useful data sets (e.g. Zheng, 2016; CSAFE, n.d.), and developing quality metrics (Nill, 2007; Yoon *et al.*, 2012; Hicklin *et al.*, 2013; Pulsifer *et al.*, 2013; Kellman *et al.*, 2014; Chugh *et al.*, 2018; Kalka *et al.*, 2020; Swofford *et al.*, 2021).

However, for probabilistic reasoning to be adopted in forensic practice, more than statistical knowledge will be necessary. Probabilistic reasoning will have to be adopted by the current workforce of forensic practitioners. It is not clear that this workforce is either knowledgeable about or committed to a probabilistic approach. Indeed, some practitioners have expressed scepticism, or downright hostility, towards probabilities and statistics (e.g. McKasson, 2001; Bush, 2009; Jayaprakash, 2013).

In addition to statistical knowledge, therefore, social scientific knowledge will be necessary to actually enact the introduction of probabilistic reasoning into forensic practice. Such knowledge can help us understand issues such as: whether and to what extent forensic practitioners understand probabilistic reasoning; how better to educate practitioners in probabilistic reasoning; and whether practitioners welcome the introduction of probabilistic reasoning or are actively resistant to it and the reasons for, and sources of, such reactions.

The present study was intended to be a contribution to that effort. It used a survey of practitioners in a single forensic discipline—friction ridge examination—and it focused on a single deployment of probabilistic reasoning, which we call 'probabilistic reporting'—that is, the reporting of forensic findings in probabilistic (as opposed to 'categorical') terms. Friction ridge analysis was chosen because the researchers are familiar with the discipline and had connections with the large practitioner community, it is a widely used and influential pattern evidence discipline, the debate over probabilistic reporting is familiar to many in the discipline, and statistical tools have been developed and are familiar to the community. This study aims to capture baseline data on reporting practices across the discipline in order to: (i) ascertain what kind of reporting language friction ridge examiners and Forensic Service Providers (FSPs) currently use, and to what extent examiners and FSPs use probabilistic reporting, (ii) gauge friction ridge examiners' attitudes towards probabilistic reporting, and the reasons for, or sources of, those reactions; and (iii) understand examiners' characterization of probabilistic reporting and what it means to report probabilistically.[1]

It is hoped that this study will be useful for scientists interested in fostering the use of probabilistic reasoning in forensic science. It may also be of interest to forensic practitioners, laboratory administrators, legal scholars, social scientists and others interested in the introduction of statistical thinking into forensic practice. The findings may help these groups better understand the degree of penetration of probabilistic reasoning that has already been achieved, the reasons practitioners may welcome or resist the introduction of probabilistic reasoning, and how to improve education and implementation efforts.

2. Background

Friction ridge impression evidence (colloquially referred to as 'fingerprint evidence') has long been considered one of the most important kinds of forensic evidence used in criminal and civil litigation

[1] The study originally had a fourth goal: to capture and record the experiences of latent print examiners who have adopted probabilistic reporting. However, we received only 6 survey responses (2%) from such examiners. We deemed this an insufficient sample, and we do not address this goal further here.

and is often regarded by jurors and other criminal justice system actors as incontrovertible proof that an individual touched an item in question (Lieberman *et al*., 2008; Garrett and Mitchell, 2013; Koehler, 2016; Ribeiro *et al*., 2019; van Straalen *et al*., 2020). This is based upon decades of testimony that fingerprint evidence is unique to an individual and that no two individuals, including identical twins, share the same arrangement of friction ridge skin (Cole, 2001). Friction ridge examination consists of visual observation and comparison of friction ridge details between two impressions. Traditionally, the process for conducting friction ridge examinations is described by the acronym ACE-V, which stands for 'Analysis', 'Comparison', 'Evaluation', and 'Verification' (Ashbaugh, 1999). ACE-V has been described in the forensic literature as a means of comparative analysis of evidence since 1959 (NRC, 2009).

For nearly a century, latent print examiners have expressed their findings in categoric terms with statements or implications of absolute certainty, something also true of many other forensic disciplines (Bali *et al*., 2020). When we characterize reporting as 'categoric', we mean that reporting follows a system in which reports are assigned to 'categories' which are treated as homogeneous within and mutually exclusive. For example, in friction ridge analysis, it is common to report results in three categories often named 'exclusion', 'inconclusive' and 'identification'. Although categoric reporting does not require statements of certainty (and often allows for statements of uncertainty), historically it has been common to treat one or both of the endpoints of categoric frameworks (i.e. 'exclusion' and 'identification' in the framework above) as statements either of certainty or of some state of quasi-certainty that can be treated as tantamount to certainty (Cole, 2014). So, e.g. lay fact-finders were often told that fingerprints 'matched' with '100% certainty' and the two impressions were made by the same source (Cole, 2007). Over time, terms such as 'match', 'identification', and 'individualization' became synonymous expressions, all of which meant that a specific individual was determined to be 'the' source of an impression (Cole, 2014). Such claims have been criticized as unsupportable by individual scientists and scholars (e.g. Robertson, 1990; Stoney, 1991; Champod and Evett, 2001; Broeders, 2006; Meuwly, 2006; Saks and Koehler, 2008; Cole, 2009; Page *et al*., 2011; Neumann *et al*., 2012; Eldridge, 2017) and a number of governmental and scientific reports (NRC, 2009; Campbell, 2011; NIST, 2012; PCAST, 2016; AAAS, 2017). While this article is not intended to review these debates, we summarize the criticisms as follows: First, statements of certainty, to the extent that they are being made, are inherently misleading and unscientific—they systematically overstate the value of the evidence. Forensic results, particularly those with an inclusionary outcome, cannot preclude the possibility of any considered hypothesis. Proper reporting of forensic results should therefore account for the probability of the evidence under the considered hypotheses and some probability, even if small, must necessarily be assigned to each hypothesis (Aitken *et al*., 2010). Secondly, statements of certainty aside, categorical frameworks are too simplistic. They treat all forensic results as equivalent, no matter how different, assigned to the same category. And, they may overstate the difference between two forensic results that are quite similar but fall on opposite sides of the arbitrarily defined boundary between two categories. Ideally, then, forensic results should be reported along a continuum rather than in categories (Champod and Evett, 2001). How this should be done is not something we will discuss in this article, but, to summarize, proposals range from expanded 'verbal scales' to expressing probabilistic statements along a continuum. Methods for expressing probabilistic findings range from the use of likelihood ratios to the use of accuracy data, and from the use of statistical models and associated software to the use of subjective probabilities based on human judgement.

The latent print discipline has responded with statements that limit strength of the claim that the words 'identification' and 'individualization' are supposed to convey (Garrett, 2009; SWGFAST, 2013). These changes, however, insisted on retaining the terms themselves and the claim that 'two impressions were made by the same source' while dispensing with the phrase '*to the exclusion of all others*', resulting in ensuing criticisms that the change had no practical impact (OSAC FRS, 2016; Cole, 2018). At least one crime laboratory, the United States Army Criminal Investigation Laboratory (USACIL), the primary forensic laboratory supporting the criminal investigative mission of the Department of Defense, announced a policy change to abandon the term 'identification' and report their findings in a probabilistic framework (Defense Forensic Science Center, 2015). In 2017, the USACIL went a step further and announced the implementation of a statistical software application, *FRStat*, to provide probabilistic support to fingerprint associations (Defense Forensic Science Center, 2017).[2] In 2018, the Organization for Scientific Area Committees (OSAC) for Forensic Science, Friction Ridge Subcommittee (OSAC FRS), which is responsible for the promulgation of standards and best practices related to the forensic examination of friction ridge skin impression evidence nationwide, released the proposed Standard for Friction Ridge Examination Conclusions (OSAC FRS, 2018), which took an additional step towards ensuring a probabilistic expression. While the proposed standard maintains the term 'identification', it was redefined in a probabilistic likelihood ratio format (OSAC FRS, 2018). In addition to the revised definition, the OSAC FRS states that 'an examiner shall not assert that a source identification is the conclusion that two impressions were made by the same source or imply an individualization to the exclusion of all other sources' (OSAC FRS, 2018).

This debate over reporting practices provides the context for the present study. However, the purpose of this study was not to advance the debate for or against probabilistic reporting. Rather, it was to try to elicit the perspectives of a practitioner community on the prospect of probabilistic reporting.

3. Methods

3.1 *Participant recruitment and survey administration*

Participants were recruited to participate in the study by invitation through their membership in the International Association for Identification (IAI), the largest professional organization of forensic fingerprint practitioners in the world, and through word of mouth by members of the friction ridge (fingerprint) community. The survey was emailed to approximately 1700 IAI members listed as having background in friction ridge examination (see Appendix A for the recruitment email). On the Study Information Sheet (but not in the Recruitment Email) participants were informed that they

[2] The *FRStat* software is method developed by the USACIL designed to serve as a quality assurance tool and a means of quantitatively conveying the significance of an association observed by an examiner. The *FRStat* development and validation is described by Swofford *et al*. (2018) Briefly described, the *FRStat* first calculates a similarity value (called GSS) between two sets of features identified by an examiner on two separate impressions which the analyst believes to correspond. The software then provides two estimates, one indicating how often prints originating from common sources would result in a GSS that is equal to or greater less than the calculated GSS and another indicating how often prints from different sources would result in a GSS that is equal to or greater than the calculated GSS. The two values are then combined as a ratio providing a single summary statistic indicating to what extent the GSS is consistent with originating from a common source compared to different sources. Generally speaking, higher values of this ratio indicate greater evidence in favour of the analyst's opinion of association; lower values indicate less evidence in favour of the analyst's opinion of association and may serve as a quality assurance tool to flag a comparison as potentially problematic due to insufficient similarity to support an association, based on the thresholds and standards set by an organization's policies.

would receive a Centre for Statistical Applications in Forensic Evidence (CSAFE)-branded coffee mug for completing the survey. Eligible participants were forensic practitioners 18 years of age or older. Participants were provided a link to an online survey using a commercial survey platform, Qualtrics®. All responses to the survey were anonymous. As will be discussed below, the survey received a total of 301 survey responses.

The survey was open for a 2-month period during August and September 2018. After giving informed consent, participants were presented with a series of questions pertaining to their demographics (gender, age and education), employment and testimony experience. Participants were then provided a closed-response question in which they were asked to choose which of three sample statements most closely resembled the wording they currently used in reports of an association between two friction ridge impressions (see Appendix B). The first option was meant to encompass a variety of different 'categoric' ways in which friction ridge examiners tend to report and testify. The second was intended to encompass the variety of ways in which friction ridge examiners currently try to testify 'probabilistically'. The third, which we call 'demonstrability', was intended to capture a kind of reporting currently advocated by some practitioners which emphasizes the 'demonstrability' of the conclusions more than their probabilistic nature (Triplett, 2016). We refer to this question as the 'trigger question' because it was used to initially divide the subject pool into two groups, labelled 'probabilistic' and 'categoric', which were administered questions slightly differently in parts of the remainder of the survey.[3] For purposes of this binary assignment, 'demonstrability' respondents were aggregated with the 'probabilistic' group.[4]

Next, all participants were given an open-response question which asked them to provide the actual language used in their written examination reports when reporting an association between two friction ridge impressions in their practice. In a follow-up binary, closed-response question, participants were then asked whether they believed their actual reporting language was 'probabilistic' or 'non-probabilistic (categorical)'.

After being divided into two groups based on their responses to the trigger question, participants were administered a series of Likert-scale and free-response questions regarding their positions towards probabilistic reporting. Likert-scale questions included five response choices indicating the extent participants agree or disagree with the statements provided. The Likert-scale response choices included: 'strongly agree', 'somewhat agree', 'neither agree or disagree', "somewhat disagree," "strongly disagree." Likert-scale responses were evaluated quantitatively and free text responses were evaluated qualitatively through researcher coding and analysis using *Atlas.ti*® software. The raw surveys and our coding are publicly available through the CSAFE data portal.[5]

[3] Participants responding to the trigger question indicating that their reporting language was 'probabilistic' were asked to report their attitudes toward probabilistic reporting: (1) *before* making a change to probabilistic reporting and (2) *after* making a change to probabilistic reporting (currently) in order to understand the degree to which their views have changed over time, if at all. This group was also asked an additional set of questions concerning what was most and least effective with helping the participant understand the importance of probabilistic reporting and gain comfort with reporting and testifying using probabilistic conclusions. However, we received only 6 survey responses (2%) from examiners who have adopted probabilistic reporting. We deemed this an insufficient sample, and we do not address this further.

[4] We doubt that either forensic statisticians or the proponents of the approach themselves would consider the statements associated with 'demonstrability' probabilistic. We combined these responses with the 'probabilistic' group because, though not probabilistic, they do represent a desire to move beyond the *conventional* categoric approach, even if the demonstrability approach still does consider itself categoric (Triplett, 2018).

[5] https://data.csafe.iastate.edu/DataPortal/#

3.2 *Current reporting practices*

3.2.1 Categoric versus probabilistic. The first aim of the survey endeavoured to capture current reporting practices for associations between friction ridge impressions. This was accomplished in two different ways. First, examiners were asked to choose from three fixed options (referred to earlier as the trigger question). Secondly, we offered participants the opportunity to articulate their reporting language in their own terms. Following this second probe, participants were asked to self-report whether they believed the language they used was 'probabilistic' or 'non-probabilistic (categorical)'. This second probe allowed us the opportunity to evaluate whether the submitted language was or was not probabilistic and compare participants' self-reports to our own evaluations. The free text responses for which participants provided samples of the actual language used in their written examination reports when reporting an association between two impressions were evaluated and coded as 'probabilistic' or 'non-probabilistic (categorical)' independently by two of the researchers using the criteria outlined below:

> 'Statements are coded "Probabilistic" if they openly and transparently assign in any way (verbal or numerical) a probability to the alternative hypothesis.
>
> Statements are coded "Categorical" if they do not assign a probability to the alternative hypothesis or if they do assign a probability to the alternative hypothesis but, in the same statement, minimize, belittle, or otherwise encourage the disregarding of that probability.'

Coding discrepancies between the two researchers were reviewed by the third researcher and discussed until a consensus was reached. This design allowed us to compare participants' self-reports to our own evaluations of whether or not statements were probabilistic.

3.2.2 Types *of categoric and probabilistic reporting.* In order to achieve greater specificity about the nature of the statements being used, we subdivided 'categoric' and 'probabilistic' statements into two subcategories each. We subdivided categoric statements into 'Traditional' or 'Elaborated', following the nomenclature proposed by Bali *et al.* (2020). Traditional statements are generally those kinds of statements that have pervaded the friction ridge discipline for the past century. Elaborated statements are those that appear to recognize that the manner of reporting needs to change, but appear to do so subtly. An example would be the redefinition of the term 'individualization' by the Scientific Working Group for Friction Ridge Analysis Study and Technology (SWGFAST) to mean 'the decision that the likelihood the impression was made by another (different) source is so remote that it is considered as a practical impossibility' (SWGFAST, 2013). More specifically, 'Traditional' statements (e.g. 'the two prints are from the same source'; 'this finger made this print'; 'the print was identified to the defendant'; 'I made an identification', etc.) assign no probability to the alternative hypothesis. 'Elaborated' statements assign a probability to the alternative hypothesis but also minimize it with a statement that encourages disregarding it (e.g. 'practical impossibility', 'negligible', 'discounted', etc.).

We subdivided probabilistic statements into two categories according to the degree of rigor with which the statements follow the logical and formal rules of probabilistic reporting (e.g. clearly

articulating hypotheses) (Evett, 2015). Statements in the first category, which we call 'Probability of Findings', tend to articulate two hypotheses and characterize the probability of the evidence. Statements in the second category, which we call 'Probability of Hypothesis', tend to articulate only one hypothesis and characterize the probability of the hypothesis (i.e. posterior probabilities) as opposed to the probability of the evidence. Based on our reading of the statement, we believe the 'Probability of Hypothesis' statements represent 'efforts' to testify in a logical probabilistic manner which have, like many efforts to speak probabilistically, inadvertently transposed the conditional (Evett, 1995).

3.2.3 Support *for probabilistic reporting (agency policy versus personal belief).* In addition to gaining a general understanding of the extent to which examiners report categorically versus probabilistically, we were also interested in understanding the extent to which examiners support reporting probabilistically. By asking examiners to report the statements that they use in their actual reports, we may have captured agency policy rather than examiners' personal beliefs. In order to better understand examiners' personal positions towards probabilistic reporting, thus capturing whether such examiners may be at least open to the idea of a transition, we asked several questions designed to elicit their personal beliefs about categorical and probabilistic reporting. These were explored using the following Likert-scale response questions for all respondents:

> 'I feel that the proposed shift away from "identification" and the use of probabilistic language is an appropriate direction for the fingerprint community.
>
> I do not understand why there is concern with expressing positive conclusions in absolute terms, such as "identification."
>
> I support probabilistic reporting because it is a scientifically more appropriate means of expressing positive fingerprint conclusions.
>
> I do not understand why probabilistic conclusions are more appropriate means of expressing positive fingerprint conclusions.
>
> I am willing to take an active role in helping other practitioners become more understanding and accepting of probabilistic reporting.'

3.3 Attitudes towards probabilistic reporting

3.3.1 Receptivity *to probabilistic reporting.* The second aim of the survey endeavoured to capture examiners' attitudes towards probabilistic reporting and the reasons for, or sources of, those reactions. This was accomplished in several ways. First, our survey question 'I feel that the proposed shift away from "identification" and the use of probabilistic language is an appropriate direction for the fingerprint community' probed the current state of examiners' receptivity to probabilistic reporting. Secondly, in order to gain greater insight into examiners' views, we solicited a free text response which invited participants to elaborate on why they agree or disagree 'that the

proposed shift away from "identification" and the use of probabilistic language is an appropriate direction for the community'. The free text responses were analysed in three groups according to aggregated responses from the Likert-scale: (1) those who perceive probabilistic reporting as *appropriate* (i.e. those who responded 'somewhat' or 'strongly agree'); (2) those who perceive probabilistic reporting as *inappropriate* (i.e. those who responded 'somewhat' or 'strongly disagree'); and (3) those who were *neutral* as to the appropriateness of probabilistic reporting (i.e. those who responded 'neither agree nor disagree'). Free text responses were single-coded by the second author in order to derive themes that emerged from the data according to a grounded theory approach. The second author has been studying friction ridge analysis from historical, sociological, epistemological and rhetorical perspectives for more than 20 years and, therefore, is, we believe, sufficiently familiar with the jargon of the discipline to interpret the responses. After a provisional list of themes was generated, the themes were re-evaluated, and some themes were aggregated, disaggregated or deleted. The researcher then made a second pass through the data using this final list of themes. There was no maximum placed on the number of themes which could be applied to any single response, but the minimum was 1 to ensure the assignment of at least one theme to each response (i.e. if a response did not fit any existing theme, a new theme was added).

3.3.2 General *opposition to probabilistic reporting*. In addition to the free text responses allowing participants to elaborate on why they agree or disagree 'that the proposed shift away from "identification" and the use of probabilistic language is an appropriate direction for the community', we were also interested in understanding key reasons for the opposition in a more structured way. We accomplished this by asking all participants several Likert-scale response questions designed to elicit general reasons why they may be opposed to probabilistic reporting. The questions were selected based on our anecdotal observations of examiners' claims or implications at conferences, online chat boards and informal discussions over the last several years:

> 'I feel that law enforcement, special agents, and/or other investigators would not understand how to interpret probabilistic conclusion language.
>
> I feel that defense attorneys would take advantage of probabilistic conclusion language to create reasonable doubt.
>
> I feel that prosecutors would be less willing to use fingerprint evidence in court because of the probabilistic conclusion language.
>
> I feel that judges and/or jurors would not understand probabilistic conclusion language.
>
> I feel that I do not sufficiently understand probabilities and would not to be able to properly testify to my conclusion in court.
>
> I feel that a probabilistic conclusion is too weak of a conclusion.
>
> I feel that a probabilistic conclusion would negatively impact the outcome of a trial.
>
> I feel that if I were to report and/or testify to probabilistic language that my certification with the IAI would be in jeopardy.

I feel that probabilistic reporting will cause the number of erroneous associations to significantly increase.'

3.4 Characterizations of probabilistic reporting

The third aim of the survey endeavoured to understand examiners' characterization of probabilistic reporting and what it means to report probabilistically. Although the concept of 'probabilistic reporting' has been advocated by proponents, it is unclear what examiners understand those words to mean and whether they differ from one another. In a free-text response question, we sought to allow the respondents to tell 'us' what they understood the term 'probabilistic reporting' to mean with the following question: 'How would you describe probabilistic reporting, compared to non-probabilistic (categorical) reporting?' For analysis purposes, we divided respondents into two groups (categoric reporters and probabilistic reporters) based on the trigger question discussed above. Using the same analysis procedures described in section 3.3.1 above, the free text responses were reviewed by one of the researchers in order to derive themes that emerged from the data according to a grounded theory approach.

4. Results

4.1 Participant responses

A total of 435 raw survey responses were received in the 2-month period the survey was open (August and September 2018); however, 134 survey responses omitted the trigger question and were largely incomplete or blank altogether, and we discarded them. Given that the survey was made freely available on the internet, many 'responses' may have reflected individuals who preferred to view the survey, rather than complete it. After removing the incomplete surveys, a total of 301 completed surveys were available for evaluation—yielding a response rate of 17.7% out of approximately 1700 IAI members invited to participate. A completed survey, however, does not mean that the respondent answered every question.

Of the 301 respondents who completed surveys, the demographics are as follows: 44% were male and 54% were female (2% unreported), 88% reported being employed in the USA and 10% reported being employed outside of the USA, 83% reported having a Bachelor's Degree or higher, 84% have testified in court, 54% testified in court in the past year, the average years of experience is 16.5 years (standard deviation of 11.2 years), and the distribution of participants' ages was: 8% reported as 20–29 years, 36% reported as 30–39 years, 28% reported as 40–49 years, 19% reported as 50–59 years, 8% reported as 60–69 years, 2% reported as 70–79 years, and 2% unreported.

4.2 Current reporting practices

4.2.1 Categoric *versus probabilistic*. Among the 301 completed surveys, all participants responded to the trigger question with three fixed reporting options. The responses to this question are shown in Table 1.

Table 1 shows that the vast majority of respondents (88%) use categoric reporting language. Probabilistic reporting is rare (10%) and only 2% of respondents use 'demonstrability' language. In short, 9 out of 10 friction ridge examiners surveyed responded that they report categorically.

TABLE 1 *Breakdown of fixed options of current reporting practices for associations.*

Response	Brief description	*n*
Categoric	Same source; identified to, matched	88% (264/301)
Probabilistic	Likelihood same or different source	10% (31/301)
Demonstrability	The conclusion is easily demonstrable to others	2% (6/301)

Examiners selected the samples of fixed reporting options that most closely resembled their own. The fixed reporting options acted as a 'trigger question' to initially categorize respondents as reporting categorically or probabilistically. The 'demonstrability' option was treated as probabilistic for purposes of this initial categorization.

TABLE 2 *Breakdown of categoric versus probabilistic reported based on respondents' self-report compared to researcher coding of the actual reporting language*

1	2	3	4	5	6
Self-report		Researcher coded		Total (Researcher coded)	
Categoric	88% (217/247)	Categoric	87% (215/247)	Categoric	98% (241/247)
		Probabilistic	<1% (2/247)		
Probabilistic	12% (30/247)	Categoric	10% (26/247)	Probabilistic	2% (6/247)
		Probabilistic	2% (4/247)		

Among the 301 completed surveys, only 247 provided free text responses with a sample of their actual reporting language. For those 247 respondents, we were able to code whether the reporting language was categoric or probabilistic (coding discrepancies between the two researchers occurred in 3 of the 247 responses—all three related to whether a categoric statement should be sub-coded as 'traditional' versus 'elaborated'). Table 2 compares respondents' self-reports with researcher coding. The self-reports for the smaller group of 247 respondents are consistent with those found in the larger group of 301 (Table 1): 88% of examiners described themselves as reporting categorically (Table 2, column 2). Surprisingly, even among the small number of examiners who purported to report probabilistically, the majority of them, in our view, report categorically. Conversely, two respondents described themselves as reporting categorically, but provided a sample statement that we interpret to be probabilistic (columns 4 and 6).

Readers may wonder whether and why they should treat our assessment of whether a statement is probabilistic as dispositive. We believe that readers who examine the statements for which we disagreed with the participant's self-report will find our assessments uncontroversial (taking note of the definitions of these two categories we provide above). For example, one respondent self-

TABLE 3 *Breakdown of categoric and probabilistic reporting into subtype based on determinations of categoric versus probabilistic from researcher coding (traditional versus elaborated for categoric and probability of findings versus probability of hypothesis for probabilistic)*

Report type	Percentage (researcher coded)	Report subtype	Percentage
Categoric	98% (241/247)	Traditional	89% (220/247)
		Elaborated	9% (21/247)
Probabilistic	2% (6/247)	Probability of findings	2% (4/247)
		Probability of hypothesis	<1% (2/247)

reported that the following statement was 'probabilistic'; however, we interpreted it to be categoric: 'Identification is the opinion of an examiner that there is sufficient quality and quantity of detail in agreement to conclude that two impressions originated from the same source.' Conversely, another respondent self-reported that the following statement was 'categoric'—we coded it as probabilistic: 'In the opinion of this examiner the likelihood that the impressions were made by a different source other than the one listed is very small.' A full listing of all the statements for which researcher coding conflicted with self-report is presented in Appendix C.

Table 2 suggests that categoric reporting is even more prevalent than the 90% figure found in the self-reports. According to our coding of actual provided sample reporting language,[6] 98% of respondents are using categoric statements to describe associations between friction ridge impressions. Only 6 of the 247 respondents provided statements that we interpret as probabilistic.

4.2.2 Types *of categoric and probabilistic reporting*. Table 3 provides the breakdown of respondents' categoric statements subdivided as 'traditional' versus 'elaborated' and respondents' probabilistic statements subdivided as 'probability of findings' versus 'probability of hypothesis'.

4.2.3 Support *for probabilistic reporting (agency policy versus personal belief)*. Table 4 shows the responses from participants who actually report categorically related to questions designed to elicit their personal support for probabilistic reporting. As Table 4 shows, between 32 and 43% of respondents who report categorically responded in ways that suggest they personally support probabilistic reporting.

4.3 Attitudes towards probabilistic reporting

4.3.1 Receptivity *to probabilistic reporting*. The first question on Table 4 shows the range of responses on the Likert-scale to the question 'I feel that the proposed shift away from

[6] Readers may wonder whether multiple participants from the same laboratory participated in the survey. The anonymous nature of the survey precludes any insights into this.

TABLE 4 *Participants' (categoric respondents based on researcher coding) responses to Likert-scale questions related to examiners' personal beliefs and support for probabilistic reporting*

Question	Likert-measure	Total	Degrees of dis/agreement aggregated
I feel that the proposed shift away from 'identification' and the use of probabilistic language is an appropriate direction for the fingerprint community	Strongly agree	9% (24/265)	32% (85/265)
	Somewhat agree	23% (61/265)	
	Neither agree or disagree	10% (26/265)	10% (26/265)
	Somewhat disagree	28% (73/265)	58% (154/265)
	Strongly disagree	31% (81/265)	
I do not understand why there is concern with expressing positive conclusions in absolute terms, such as 'identification'	Strongly agree	22% (59/265)	49% (129/265)
	Somewhat agree	26% (70/265)	
	Neither agree or disagree	9% (23/265)	9% (23/265)
	Somewhat disagree	25% (65/265)	43% (113/265)
	Strongly disagree	18% (48/265)	
I support probabilistic reporting because it is a scientifically more appropriate means of expressing positive fingerprint conclusions	Strongly agree	12% (34/280)	37% (104/280)
	Somewhat agree	25% (70/280)	
	Neither agree or disagree	12% (33/280)	12% (33/280)
	Somewhat disagree	21% (58/280)	51% (143/280)
	Strongly disagree	30% (85/280)	
I do not understand why probabilistic conclusions are more appropriate means of expressing positive fingerprint conclusions	Strongly agree	21% (59/280)	46% (128/280)
	Somewhat agree	25% (69/280)	
	Neither agree or disagree	14% (38/280)	14% (38/280)
	Somewhat disagree	25% (69/280)	41% (114/280)
	Strongly disagree	16% (45/280)	
	Strongly agree	13%	34%

(*Continued*)

Table 4 *(continued)*

Question	Likert-measure	Total	Degrees of dis/agreement aggregated
I am willing to take an active role in helping other practitioners become more understanding and accepting of probabilistic reporting	Somewhat agree	(37/280) 20% (57/280)	(94/280)
	Neither agree or	34% (95/280)	
	Somewhat disagree	13% (35/280)	33% (91/280)
Strongly disagree	20% (56/280)		

"identification" and the use of probabilistic language is an appropriate direction for the fingerprint community' for respondents and then shows aggregations of the Likert-scale responses into broader categories. From these data, we see most friction ridge examiners feel that probabilistic reporting is not an appropriate direction for the community (58%). Few examiners are neutral on the issue (10%). Among the 239 respondents with opinions (responding 'somewhat' or 'strongly dis/agree' in either direction), approximately two-thirds view it as inappropriate and one-third view it as appropriate. Although the proportion who view it as appropriate is far greater than those who actually apply probabilistic reporting, the majority of the examiner community at large still remains generally opposed to, or at least sceptical of, probabilistic reporting.

When invited to elaborate on why participants agree or disagree 'that the proposed shift away from "identification" and the use of probabilistic language is an appropriate direction for the community' in a free text response, respondents expressed a wide diversity of opinions to this question. Some respondents wrote long disquisitions, and at least one complained about the lack of space in which to enter a response. The mean and median number of themes for each response was 3. The minimum was 1 (our coding rules required the assignment of at least one theme to each response), and the maximum was 8.

4.3.1.1 Respondents supporting probabilistic reporting Among the 85 participants who consider probabilistic reporting 'appropriate' (i.e. those who responded 'somewhat' or 'strongly agree' on the first question in Table 4), a total of 36 different themes were identified suggesting 'why' they viewed it as appropriate, and a total of 113 themes were coded across the 85 respondents. Table 5 shows all themes that were mentioned by more than one respondent. The full list of themes is available through the CSAFE data portal (see footnote 5).

The most common responses were that probabilistic reporting was an improvement over past or current practice. This was most commonly described as either 'more accurate' or 'more scientific'. An example of a 'more accurate' statement is:

TABLE 5 *'Shift to probabilistic reporting is appropriate' : coded themes mentioned more than once (n = 85)*

Number	Theme	Frequency
1.	More accurate	12
2.	More scientific	12
3.	Uncertainty	10
4.	Jury clarity	9
5.	Transparency	7
6.	Weight of evidence	7
7.	Finer	6
8.	Uniqueness unproven	6
9.	Objectivity	5
10.	It's happening	3
11.	Statistic ok with verbal	3
12.	Appropriate	2
13.	Consistent with other disciplines	2
14.	External scrutiny	2
15.	Law	2
16.	Overselling	2
17.	Reliance on stock phrases/expertise	2
18.	Sound reasoning	2
19.	Step in right direction	2

> 'It is a more accurate description of my observations and the limits of my observations' (38).[7]

An example of a 'more scientific' statement is:

> 'I think it's time for LP [latent print] examination to apply more scientific rigor to the practice of latent print examination. This would include articulating probabilities when reporting results of examinations. I would hope that this ultimately leads to more credibility for friction ridge examination as a forensic discipline' (416).

These respondents also viewed probabilistic reporting's ability to convey uncertainty as an advantage. For example, one respondent wrote:

> 'We shouldn't speak in absolute terms. Identification over estimates the evidence and our conclusions should convey the level of certain [sic] we know. Even if we know what we mean by identification it does imply to a jury absolute certainty. It's not scientific. I would like the field to be better (31)'.

[7] All quotations or references to specific responses from the survey are followed by a parenthetical reference to the respondent number. Respondent numbers were assigned to all respondents who opened the survey, including those who did not complete it. Therefore, some respondent numbers are greater than the total number of respondents, 301. Spelling errors in the responses are corrected, but grammatical errors are not.

Several respondents specifically cited the 'weight of evidence', an important concept in forensic statistics, as an advantage of probabilistic reporting. Many of these respondents perceived probabilistic reporting as offering greater 'jury clarity' and 'transparency' for the jury. This contrasts with many respondents who do not support probabilistic reporting and cited 'jury confusion' as a reason for their opposition (see below).

Overall, respondents offered a rich and diverse set of reasons for the appropriateness of probabilistic reporting. There was reference to epistemological considerations (e.g. 'uniqueness unproven'; 'objectivity'), external forces (e.g. 'it's happening', 'external scrutiny', 'law'), and perceived problems with current practice (e.g. 'overselling', 'reliance on stock phrases/expertise', 'conclusions are not overstated'). However, in contrast to the responses that probabilistic reporting is inappropriate, those who consider it appropriate were characterized by a clearly perceptible degree of ambivalence. Many responses coded as 'appropriate' listed some advantages of probabilistic reporting, but then turned to some perceived disadvantage.[8] An example is this:

> 'It gives a more accurate representation of the validity of the conclusion and results reported. However, I do worry it may confuse the issue in the event of a distracted/inattentive jury (419)'.

Recognizing the importance of this ambivalence, we coded such responses as containing 'reservations'. Fully 36 of the 85 responses (42%) that considered probabilistic reporting appropriate contained such reservations. Such ambivalence was not nearly as common among the responses that considered it inappropriate, and so 'reservations' in those responses were not counted.

4.3.1.2 Respondents not supporting probabilistic reporting Among the 154 participants who considered probabilistic reporting 'inappropriate' (i.e. those who responded 'somewhat' or 'strongly disagree' on the first question in Table 4), a total of 49 different reasons were identified suggesting 'why' they viewed it as inappropriate, and a total of 265 themes were coded across the 154 respondents. Table 6 shows all themes that were mentioned by more than one respondent. The full list of themes is available through the CSAFE data portal (see footnote 5).

The most common response was that probabilistic reporting would confuse the jury:

> 'Our job in court is to make the jury understand the evidence. I feel that the language being referred to will just confuse a lay person (36)'.

For many examiners, the 'confusion' they feared lay in the move away from communicating results in the simple form of certainty:

[8] It is important to recall that in this analysis respondents were categorized according to their responses to the Likert-scale question, 'not' according to researcher coding of their free-text responses. Thus, a respondent who reported that they 'somewhat agree' that probabilistic reporting is appropriate and submitted a free-text response criticizing probabilistic reporting was still analysed in the 'appropriate' group. For example, the following free-text responses were made by respondents who self-reported that they agreed or somewhat agreed that probabilistic reporting was appropriate: 'I believe much more study into the usage of probabilistic statements is required. Specifically, in district and county courts within the US, not just military court as is currently being done' (86); and 'The shift towards probabilistic language seems contradictory to the principles of fingerprint identification that I learned in my training years. I was taught to give definitive conclusions—not maybes. Additionally, I worry that probabilistic conclusions may have the negative effect of increasing the number of people being erroneously associated with a given latent print' (374).

TABLE 6 *'Shift to probabilistic reporting is inappropriate': coded themes mentioned more than once (n = 154)*

Number	Theme	Frequency
1.	Jury confusion	36
2.	Probability not ready	26
3.	Underselling	23
4.	Quantification impossible	18
5.	Opinionization	15
6.	Uncertainty can be eliminated	14
7.	Misleading	12
8.	Uniqueness	12
9.	Wealth of empirical data	9
10.	DNA paradigm inappropriate	7
11.	Unnecessary	7
12.	Appropriate in some cases	6
13.	Combine probability with id	6
14.	Models do not capture all information	6
15.	Vague	6
16.	Only problem uncertainty	5
17.	Verbal/subjective probabilities unacceptable	5
18.	Politics	4
19.	Sceptical of statistics as discipline/all models wrong	4
20.	Customer is police/attorneys	3
21.	Examiner confusion	3
22.	Risk contradicting ground truth more often	3
23.	Transparency sufficient	3
24.	Whole population problem	3
25.	Accuracy more important	2
26.	Defence exploitation	2
27.	Does not prefer	2
28.	Not science	2

> 'This shift seems to add confusion. It's giving some degree of uncertainty to our conclusions (185)'.

This contrasts with the respondents favouring probabilistic reporting who viewed 'uncertainty' as an advantage.

A similar concern was invoked when examiners expressed concern about 'underselling'—the concern that probabilistic reporting is too 'weak' and would understate the probative value of the evidence:

> 'No accurate or "full" way to express LPE's opinions. The numbers are weak and meaningless (413)'.

Compared to respondents who consider probabilistic reporting appropriate, some examiners cited 'jury confusion' to express the opposite concern: that probabilistic reporting would overstate, rather than understate, the value of the evidence:

> 'Until probabilistic language is valid and reliable, as well as easily understood by all practitioners and easily explained by those practitioners to a judge and jury, they are useless numbers. They provide to confuse the trier of fact, add little to the data, and very likely will serve to bolster the testimony and evidence (24)'.

The second most common reason given was 'probability not ready'. These respondents communicated that they were not opposed to probabilistic reporting in principle. Their opposition, rather, was a practical consideration based on their perception of the current state of affairs with regard to the development of a statistical model useable for assigning a probability to a friction ridge association:

> 'There is no current scientific basis for a probabilistic model. Reporting latent print conclusions in probabilistic language is misleading and unscientific. Statistics are not in themselves scientific or objective (254)'.

Some other common themes cited fundamental epistemological concerns about the use probability to communicate the findings of friction ridge analysis. One such concern is 'quantification is impossible':

> 'I believe there cannot be probabilistic language involved with latent print examination. There are too many variables involved. Latent prints examination is too subjective to have probability (169).'

These responses suggest the impossibility of quantification as deriving from the limits of friction ridge analysis: it is inherently uncertain, subjective, and reliant on continuous rather than discrete measures and, therefore, impervious to quantification. In contrast, other respondents resisted the premise that friction ridge analysis is inherently uncertain, claiming 'uncertainty can be eliminated':

> 'When I form the opinion that I have made an identification - I am certain that this is the person. There is no probability of it being from someone else. If I am not certain then I will not say it's an identification but that I can not exclude them as being the contributor (258)'.

> 'Ident is ident I don't see the purpose of assigning a probability number, it's 100% or I wouldn't call it a ident (172)'.

A third such concern is reference to 'uniqueness':

> 'Strongly disagree because it is impossible to apply a probability to something, anything, that is unique. The scientific basis in biology and other natural sciences is well established. Therefore, probabilistic language is inappropriate and unscientific (270)'.

TABLE 7 *'Neutral on shift to probabilistic reporting is appropriate' : coded themes mentioned more than once (n = 26)*

Number	Theme	Frequency
1.	Jury confusion	5
2.	Truly undecided	4
3.	Does not understand	2
4.	Unnecessary	2

Some respondents expressed outright hostility towards the discipline of statistics, which was called, e.g. a 'bandwagon' and a 'fad' (239). Another respondent commented:

> 'The probabilistic method (not language) is put upon us because it is supposed to be more scientific. It comes from the DNA science. I see that nowadays DNA evidence in practice is accepted as Empirical fact in spite of statistical humdrum (296)'.

There was also some hostility expressed towards DNA, primarily for imposing an inappropriate paradigm on friction ridge analysis. 'Politics', defence attorneys, defence experts and 'critics' were also perceived as imposing the shift towards probabilistic reasoning on the discipline.

4.3.1.3 Respondents neutral to probabilistic reporting Among the 26 participants who were 'neutral' as to the appropriateness of probabilistic reporting (i.e. those who responded 'neither agree nor disagree' on the first question in Table 4), a total of 27 different reasons were identified, and a total of 36 themes were coded across the 26 respondents. Table 7 shows all themes that were mentioned by more than one respondent. The full list of themes is available through the CSAFE data portal (see footnote 5).

4.3.2 General *opposition to probabilistic reporting*. Table 8 shows the responses to the Likert-scale questions designed to elicit key reasons why examiners may be opposed to probabilistic reporting and provides them in rank order based on the proportion of examiners who responded 'agree' or 'strongly agree' to each question.

4.4 Characterizations of probabilistic reporting

When asked to describe probabilistic reporting compared to non-probabilistic (categorical) reporting in a free text response, 192 participants responded. Among those, 177 were respondents who had self-reported that they report categorically and 15 were respondents who had self-reported that they report probabilistically. Due to the small number of respondents indicating they report probabilistically, we only discuss the results of the 177 participants who indicated they report categorically. Among the 177 categoric respondents, a total of 94 different themes were identified, and a total of 326 themes were coded across the 177 respondents. Table 9 shows all themes that were mentioned by more than one respondent. The full list of themes is available through the CSAFE data portal (see footnote 5).

The most common response characterized probabilistic reporting as the quantification, or at least the communication, of 'uncertainty':

TABLE 8 *Participants' (categoric respondents based on researcher coding) responses to Likert-scale questions related to possible reasons for opposition to probabilistic reporting (from Tables 6 and 7) in rank order*

Key reasons for opposition to probabilistic reporting (in rank order)	Agree or strongly agree, % (*n*/*N*)
Defence attorneys would take advantage of probabilistic conclusion language to create reasonable doubt	79 (211/266)
Judges and/or jurors would not understand probabilistic conclusion language	79 (209/264)
Law enforcement, special agents, and other investigators would not know how to interpret probabilistic conclusion language	69 (184/265)
A probabilistic conclusion is too weak of a conclusion	48 (127/265)
I do not sufficiently understand probabilities and would not be able to properly testify	44 (116/266)
A probabilistic conclusion would negatively impact the outcome of a trial	41 (110/266)
Prosecutors would be less willing to use fingerprint evidence in court	38 (100/266)
Probabilistic reporting will cause the number of erroneous associations to significantly increase	28 (80/281)
My certification with the International Association for Identification would be in jeopardy	15 (40/266)

Note: n's vary by question.

> 'Probabilistic reporting assigns uncertainty to each examination while non-probabilistic reporting offers a conclusion (16)'.

The second most common response characterized probabilistic reporting in terms of 'quantification':

> 'Using numbers, statistics, and/or frequencies to explain a conclusion rather than words or descriptions (181)'.

As shown in Table 9, two of the more common themes that appeared in response to questions about the appropriateness of probabilistic reporting also appeared in response to questions about its definition: 'jury confusion' and 'probability not ready'.

A number of specific technical statistical concepts appeared in the responses. The notion of 'random match probability' appeared more frequently than 'weight of evidence'. This is notable because the latter of the two terms better captures the current thinking among forensic statisticians, especially with regard to pattern evidence, such as friction ridge analysis. For example, many respondents described probabilistic reporting in terms of a random match probability:

> 'I would describe it as using a statistic [sic] model to convey the likelihood of finding someone else with the same characteristics in those prints (25).'

Slightly fewer described the weight of evidence:

> 'Probabilistic reporting would involve some sort of calculation to give a weight to the conclusion based on the information present in the known and unknown (180).'

TABLE 9 *Describe probabilistic reporting—Categoric respondents: coded themes mentioned more than once (n = 177)*

Number	Theme	Frequency
1.	Uncertainty	25
2.	Quantification	19
3.	Jury confusion	17
4.	Probability not ready	17
5.	Random match probability	16
6.	Weight of evidence	15
7.	Likelihood ratio inverted conditional	13
8.	Scientific	11
9.	Likelihood ratio	9
10.	Confusing for customer	7
11.	Undermining of fingerprinting	7
12.	Continuum	6
13.	Incomplete	6
14.	Vague	6
15.	Confusing	5
16.	Confusing to practitioners	5
17.	Do not know	5
18.	Sliding scale	5
19.	Too weak	5
20.	Underselling	5
21.	Unnecessary	5
22.	DNA model	4
23.	Driven by critics/self-serving agenda	4
24.	Error prone	4
25.	Statistical model	4
26.	The way forward	4
27.	Disaster	3
28.	Misleading	3
29.	More data to support conclusion	3
30.	Possible associations	3
31.	Problematic	3
32.	Anti-ground truth	2
33.	Consider with other evidence	2
34.	Dropping exclusion of all others	2
35.	Expectation of seeing similarity	2
36.	Greater exploitation of evidence	2
37.	Imprecise	2
38.	Inappropriate	2
39.	Insufficient weight	2
40.	More accurate	2
41.	Not opposite of categorical	2
42.	Objective	2
43.	Probability impossible	2

(*Continued*)

Table 9 *(continued)*

Number	Theme	Frequency
44.	Protects incompetence	2
45.	Reliance on technology	2
46.	Score-based	2
47.	Transparent	2
48.	Wiggle room for witnesses	2
49.	Wordy	2
50.	World population paradigm	2

The likelihood ratio, which is currently a very common topic of discussion in forensic statistics, was mentioned more frequently than either random match probability or weight of evidence. However, it was more common to describe the likelihood ratio incorrectly, with the conditional inverted (Evett, 1995), than it was to describe it correctly. For example, one respondent correctly characterized the likelihood ratio:

> 'Probabilistic reporting puts a number on the result. In some models, that number is a similarity score, sort of like the scores we get when we search a print in AFIS. In other models, the number is a likelihood ratio, telling you how likely it is that the prints having so many features in common come from the same source versus how likely it is that prints having those features in common come from a different source (20).'

But others described it with the conditional inverted:

> 'Instead of just saying "identification" which would be categorical reporting, probabilistic reporting would include the likelihood of the latent print being made by the subject (189).'

5. Discussion

5.1 Participant responses

The usable response rate of 17.7% (301 out of approximately 1700 IAI members invited to participate) poses a limitation on our survey. Viewed in combination with the recruitment scheme and voluntary participation, it is difficult to know whether these responses were representative of the views of the latent print examiner population at large. However, the demographics of our participants hint that our participants were not an unusual subset of the IAI membership. Most of our participants were mid-career latent print examiners between the ages of 30 and 50 years with testimony experience.

5.2 Current reporting practices

Probabilistic reporting appears to remain rare in the friction ridge discipline, despite its adoption by a small number of FSPs and practitioners. Approximately 98% of friction ridge examiners surveyed

report categorically using terms and phrases that are reminiscent of over a century of practice. This is consistent with the findings of other studies in a variety of disciplines. In their analysis of mock forensic reports from proficiency tests, Bali *et al.* (2020) found categorical statements in 100% of toolmark reports, 100% of fibre analysis reports, 98% of firearms examination reports, 97% of glass analysis reports, 93% of questioned documents reports, 87% of handwriting examination reports, 85% of paint analysis reports and 79% of shoeprint impression reports. Across all disciplines, 94% of reports were categorical.[9] Morrison *et al.* (2016) found that categorical reporting was by far the most common for speaker identification results. One interesting observation from our survey is that approximately 80% of the examiners surveyed who claimed to be reporting probabilistically gave as examples statements that were actually categorical. This suggests some examiners may have a false belief that they are reporting probabilistically and therefore may not recognize many of the concerns over categoric reporting are applicable to them.

Even among those who report categorically, fewer than 10% appear to have adopted the 'elaborated' approach espoused by SWGFAST as early as 2011. Nearly, a decade of calls by the scientific community to move towards probabilistic reporting seem to have had limited impact. However, as noted above, approximately one-third of respondents report categorically but responded in ways that suggest they personally support probabilistic reporting. This indicates that some examiners may be 'captured' by agency policy—reporting in a manner dictated by agency policy rather than by personal belief. This raises several open questions. Why is there such a large discrepancy between examiners who support the idea of probabilistic reporting and those who actually practice it? Are there some examiners whose reporting is inconsistent with what they personally believe is appropriate? Or did these responses merely reflect the respondents' perception that probabilistic reporting was the 'socially desirable' answer?

5.3 Attitudes towards probabilistic reporting

Most respondents perceived probabilistic reporting as inappropriate for friction ridge analysis. Only around one-third of our respondents described a shift towards probabilistic reporting as appropriate. At the same time, it can reasonably be argued that finding approximately one-third of friction ridge practitioners described probabilistic reporting as 'appropriate' would have been unthinkable as recently as one decade ago, let alone two or three decades.

Longstanding 'myths' about friction ridge evidence—for example, the claim that the 'uniqueness' of friction ridge skin eliminates uncertainty in associations between friction ridge impressions—did appear in our data, and so they cannot be considered 'dead' and still lurk behind the scenes (Cole, 2014). It should be emphasized, however, that they were uncommon among our survey respondent population, who—precisely because they took the time to complete the survey—may reasonably be presumed to more aware of, and interested in, current debates and developments within the discipline.

In assessing the reasons respondents offered for their attitudes, it may be helpful to distinguish between what we might call 'consumption' issues—relating to how evidence is used by other criminal justice system actors—and what we might call 'technical' issues concerning the merits of probabilistic reporting itself. Among the majority of respondents who viewed the shift towards probabilistic

[9] It should be noted, however, that Bali *et al.* coded categorical statements 'absent' or 'present'; thus, a single forensic report could contain both a categorical and a probabilistic statement. We, in contrast, coded 'categorical' and 'probabilistic' mutually exclusively: a single statement could not be both probabilistic and categorical.

reporting as inappropriate, the degree of concern about 'consumption' issues—fact-finder comprehension, prosecutor interests, and defence exploitation—is conspicuous. For example, from the data in Table 8, we see that the primary reason respondents opposed probabilistic reporting was the concern that defence attorneys will take advantage of the probabilistic conclusion to sow reasonable doubt (79%). Respondents further supported this through a number of free text responses, such as:

> 'I believe it will confuse the jury and give the defense a chance to place reasonable doubt (425).
>
> I don't know why we would be testifying to 'probable' outcomes—it would make cross examination significantly more difficult for the expert witness (310).
>
> The conclusion that would essentially replace "identification", in my opinion, could easily be misunderstood by jurors as meaning that there is reasonable doubt in the "same source" conclusion (409).'

This was closely followed by the concerns that judges, jurors, law enforcement and other investigators would not understand or know how to interpret the probabilistic conclusion language (79% and 69%, respectively). Similarly, the leading free-text reason for probabilistic reporting being inappropriate was 'jury confusion', a related 'consumption' issue that continues to be the subject of on-going research (Garrett and Mitchell, 2013; Langenburg *et al*., 2013; Thompson and Newman, 2015; Bayer *et al*., 2016; Garrett *et al*., 2018; Thompson *et al*., 2018; Eldridge, 2019). When considering various options for reporting, it seems reasonable for examiners to express concern whether consumers of those reports are able to take appropriate action based on the information. However, whether the concerns are warranted depend on how 'confusion' is being defined. Several of the free-text responses suggest those respondents who view probabilistic reporting as inappropriate on the basis of 'juror confusion' do so over claims that jurors want a binary answer of 'yes' or 'no', probabilistic language seems 'wishy-washy', and expressing a probabilistic view does not align with the examiner's belief that the two impressions were made by the same source. For example, some respondents stated:

> 'Why would we be needed if we used the term probabilistic. That means probably him or her. SO if we say that then there is no need for fingerprint examiners. Then I guess you could say he probably wasn't there (199).
>
> I think the matter is being overcomplicated and we are watering down our testimony, which is a disservice to the victims (56).
>
> Meaningless to jury, not accurate, wishy-washy (222).'

This leads us to question whether 'juror confusion' is being used as another way of expressing concern that jurors might not place as much weight on the conclusion as the examiner believes they should (compared to a traditional categoric identification conclusion).[10] For example, one respondent stated:

[10] Although, it is unclear how much weight 'should' be placed on the conclusion given traditional categoric statements of single-source attribution have been criticized as unsupportable by individual scientists (e.g. Robertson, 1990; Stoney, 1991; Champod and Evett; 2001; Broeders, 2006,;Meuwly, 2006; Saks and Koehler, 2008; Cole, 2009; Page et al., 2011;

> 'Probabilistic Identification is Deceitful. The purpose of having Expert witness is so they can state their belief; not force the blame on to the Jury (334).'

The respondent's invocation of the notion of 'blame' is interesting. It recalls Wells's (1992) remark that fingerprint examiners' testimony that their conclusion the person of interest is the source of an unknown impression may be particularly persuasive to fact-finders because it satisfies a 'bidirectional test': if the ultimate fact is wrong, the evidence must be wrong as well. As Wells notes, this mode of reporting evidence shifts the moral hazard of legal decision-making from the fact-finder to the expert. If the conclusion is wrong, the fact-finder can reason that they were misled by the expert. Wells found that fact-finders prefer such evidence even to statistically equivalent evidence in which the moral hazard is not assumed by the expert—the expert simply states the evidence, rather than the probability of the ultimate fact (i.e. the posterior probability). The respondent above was not merely willing to assume the moral hazard of the evidence; they appeared to perceive it as 'deceitful' 'not' to. This position is in marked contrast to a common argument in forensic science which insists that experts should report only the evidence (ENFSI, 2015). However, it is consistent with our anecdotal impressions that many forensic experts genuinely believe that 'the legal system' prefers, or even requires, them to state their beliefs about posterior probabilities, rather than confining their reports to just the evidence. This may reflect a belief that the expert's posteriors are superior to the fact-finder's, or it may simply reflect a sense of obligation.

Put in simple terms, respondents were less concerned that probabilistic reporting was 'wrong'—although there was certainly a significant number of respondents who espoused that view—than they were that defence attorneys would take advantage of uncertainty or that it would mislead, or be misunderstood by, other criminal justice system actors, such as jurors, judges, attorneys, and police investigators. Arguably, such concerns are external to the discipline of friction ridge analysis and belong the realm of policy, rather than science. Admittedly, this is a complicated issue and raises questions as to the extent institutional factors could be at odds with scientific advancements for the forensic sciences. It could be argued that this is a second-order problem and one that the forensic scientist need not necessarily face alone. Instead, the "consumption problem" can reasonably be construed as falling within the purview not solely of forensic scientists, but perhaps also of lawyers and legal scholars, social scientists, educators, and policy-makers. Although it is clearly a barrier to probabilistic reporting, whether it 'should' be a barrier to probabilistic reporting and how to mediate practitioners' concerns over these issues remain open questions.

Also of interest is the concern that 'probabilistic reporting will cause the number of erroneous associations to significantly increase' shared by just over a quarter of the practitioner community. This refers to an issue that has long lurked behind proposals for probabilistic reporting: that it would enable the use of 'more' friction ridge evidence, not less, albeit evidence of more marginal value (Champod and Evett, 2001). In this sense, respondents who agreed with this prompt may have been expressing anxiety about the transition from the comforts of the categorical regime to the discomforts of the probabilistic. Under the categoric regime, experts could feel reassured by the belief that the categorical regime was ostensibly conservative: evidence whose strength was less than overwhelming was ruthlessly discarded, and thus it was believed that it was unlikely that comparisons that fell into the 'identification' category would result in error. In a probabilistic regime, however, all evidence, no matter how marginal, can in theory be reported. While these reports will be 'truer'

Neumann et al., 2012; Eldridge, 2017) and a number of governmental and scientific reports (NRC, 2009; Campbell, 2011; NIST, 2012; PCAST, 2016; AAAS, 2017).

to the weight of the evidence and potentially provide more information for investigators and courts to consider, if fact-finders interpret them as equivalent to the old categoric terms (e.g. 'identification',) they could interpret the evidence as having 'more' weight than was intended by the examiner, thus creating an opportunity for fact-finders to, e.g. improperly infer that a person of interest is 'the' source of an impression. Indeed, some free-text responses also touched on what we interpret to be this concern:

> 'I think it creates room for errors (326)
>
> I am concerned with using probabilities because I think that it will wrongfully convict someone based on a % of something that I wouldn't consider reliable as a print examiner (179).
>
> I believe probabilistic language will confuse and mislead the jury. I think there will be a spike in wrongful convictions (393).'

In contrast to the 'consumption' issues discussed above, only 15% of examiners reported that they were motivated by concerns that their certification would be placed in jeopardy if they were to report probabilistically. Feedback on this source of opposition was solicited to evaluate the influence of the historical policy of the IAI that codified longstanding culture at the time and formally remained in effect for over 30 years. Between 1979 and 2010, the IAI officially opposed any testimony or reporting of 'possible, probable, or likely friction ridge identification' with the threat of formal charges of 'conduct unbecoming' and revocation of professional certification (IAI, 1979; 1980; 2010). This formal opposition to probabilistic reporting has undoubtedly shaped the perspectives of many experienced practitioners. However, it appears to be less important to our respondents than the 'consumption' issues.

5.4 Characterizations of probabilistic reporting

Our findings suggest that examiners' characterizations of probabilistic reporting are quite diverse. A review of Table 9 suggests that survey participants fell into two broad categories in responding to this question. A significant number responded in what we might call 'technical' terms. They responded by describing back to us what, technically, probabilistic reporting is: quantification, uncertainty, 'weight of evidence' and so on. Some used specific types of measures, such as random match probability and likelihood ratio. Given the current discourse around forensic interpretation in, e.g. this journal, it is interesting that 'Bayes' Theorem' was only mentioned once.

A second group responded to the question by treating 'probabilistic reporting' as an institutional phenomenon. They described back to us what the move towards probabilistic represented within the context of institutional debates over the role of forensic science in the criminal justice system. This elicited a number of what we might call 'skeptical', and at times even cynical responses, e.g. 'undermining of fingerprinting', 'unnecessary', 'driven by critics/self-serving agenda', 'disaster', 'misleading', 'problematic', 'imprecise', 'protects incompetence', 'wiggle room for witnesses', 'wordy', 'baseless', 'guessing', 'hysteria', 'massive undertaking', 'meaningless', 'not a panacea for error', 'quantification for its own sake', 'scary to most examiners', 'scientific veneer', 'something we don't do', 'threatening to examiners careers', 'uncharted territory' and 'unhelpful'. One respondent

described probabilistic reporting simply by: 'Mt. Everest–we are going to lose many'. While we admit that we are not entirely certain how to interpret that remark, it captures the general tone of anxiety that pervaded many responses well enough that we feel it makes an appropriate title for this article.

However, the same prompt also elicited many positive responses, e.g. 'the way forward', 'transparent', 'appropriate', 'natural evolution of science', 'safer' and 'tool'. In retrospect, we can see that the open-endedness of this prompt allowed respondents to choose whether to respond in technical or non-technical terms and, to some extent, invited them to editorialize. As with the responses to our other open-ended questions, one important observation is the heterogeneity of perspectives within the friction ridge discipline—not only as it relates to what it means technically to report probabilistically, but also how examiners have characterized it as a concept. These findings are important as they suggest that even among those who might be welcoming of probabilistic reporting, there are many different perspectives as to what it means and how it might be accomplished. We note, however, that it is possible respondents could have been primed by previous survey questions, so we cannot be sure that these responses truly reflect the respondents' unadulterated opinions.

6. Conclusion

The purpose of this survey was to provide proponents of probabilistic reporting with a sense of the state of progress in one important forensic discipline: friction ridge examination. We found that probabilistic reporting has not been widely adopted and remains extremely rare. Among those who responded to our survey, 98% of respondents continue to report categorically with explicit or implicit statements of certainty. Although we found that approximately one-third of respondents evinced receptivity to probabilistic reporting, which may well represent a more receptive audience than some might have expected, we also found that significant resistance to probabilistic reporting remains across the discipline.

The most common reasons for opposition to probabilistic reporting, shared by approximately 80% of respondents, were that defence attorneys would take advantage of the uncertainties as a litigation strategy and that probabilistic language would mislead, or be misunderstood by, other criminal justice system actors, such as jurors, judges, attorneys and police investigators, respectively. Free-text responses related to their opposition were diverse and not limited to issues of whether probabilistic reporting is scientifically more appropriate. In fact, some respondents acknowledged probabilistic reporting may be more scientifically appropriate yet continued to defend traditional categoric reporting practices. Rather, attitudes towards probabilistic reporting appear to be influenced by educational, philosophical, psychological and complex judicial implications and long-standing cultural and institutional norms.

For forensic statisticians looking for guidance, we believe our findings offer three useful lessons. First, we would emphasize the sheer heterogeneity of the responses found in, e.g. Tables 5–7. Practitioners' perspectives, even on a narrowly framed issue such as probabilistic reporting for a single forensic discipline, are quite varied and complex. This will present a challenge to educators and trainers. They will not face a handful of widely held 'myths' that they need to debunk or perspectives that they need to realign. Instead, they will face a diverse array of strongly held opinions about what, if anything, ails the friction ridge discipline, how it can and should be improved, and to what extent statistics offers a solution to those problems or

would be the cause for other problems. Probabilistic reporting in latent print examination is not a 'bi-partisan' issue; it is more complicated than that.

Secondly, we would direct readers' attention to our finding that, what we have called 'consumption' issues, seem to dominate respondents' attitudes towards probabilistic reporting. This suggests something important about where our respondents perceive the boundaries of their role as experts to lie. Our respondents appear to believe that as experts, they are responsible not only for the evaluation and articulation of the evidence, but also for how that evidence will be consumed by litigators and the decisions about the evidence that 'they believe' fact-finders 'should' be making.

We believe the appropriate role of the expert is narrower. We believe that litigation strategies and juror concerns are not within the remit of the forensic scientist to do forensic science properly. Rather, we believe the role of the expert is to educate the fact-finder about the evidence and report their findings within the limits of what the science can support, but leave it to litigators to fit that information into their arguments and to fact-finders to weigh that information when making their ultimate decisions. This requires experts to neutrally represent the evidence and clearly articulate the strengths and limitations related to those findings so that fact-finders can make an informed decision, but resist the temptation to 'simplify' the evidence for fact-finder consumption (especially when such simplifying would entail rounding the probative value of the evidence up, as when a strong belief that two impressions derive from the same source is expressed as 'the impressions originate from the same source'). To be sure, we are not denying the importance of fact-finder comprehension of statistical evidence, which is well understood to be an important problem and is the subject of a wealth of research. However, we are surprised at the degree to which bench practitioners seem to understand it be 'their' problem as opposed to a problem for legal actors, psychologists, policy-makers, etc. This puts statisticians in a bind because many practitioners seem to view probabilistic reporting inappropriate not because it is an incorrect way to report evidence, but rather because fact-finders have difficulty in understanding statistics. The latter point is undoubtedly true, but it also may well be an intractable problem. In this sense, concerns about consumption can begin to seem like stalling tactics.

From our findings, few of the respondents appear to share this view. This is most evident in how they approached several of the open-ended questions. When asked about the scientific merits of probabilistic reporting over categoric reporting or how they characterize probabilistic, compared to categoric, reporting they responded by discussing how such reporting language could be (mis)used in court. This presents another challenge. If experts are expected to report their findings within appropriate limits, the role of the expert will need to be clarified by policy-makers and enforced by the judiciary. As long as experts are allowed to express their opinions categorically, they will continue to do so. Proponents of probabilistic reporting, therefore, will need to not only include forensic statisticians, educators and trainers devising statistical tools and recommending reporting frameworks for experts, but also policy-makers and members of the judiciary to require it of experts and enforce it during litigation.

Thirdly, our survey suggests that many respondents do not share a common understanding of what is meant by the term 'probabilistic reporting'. We would point to the high number of respondents who claimed to be reporting probabilistically but, in fact, were not. The survey results offer forensic statisticians ample further reasons for pessimism about their educational efforts thus far: the insistence that uncertainty can be eliminated in friction ridge analysis, the claim that quantification of friction ridges is impossible, the claim that a statistical model is 'impossible' for the same reason claims of certainty are impossible, and the scepticism and mistrust directed towards statistics as a

discipline. We might even go a step further and suggest that many respondents did not understand what it means to report probabilistically or why categoric reporting, as we define it, has come under so much criticism. However, this was not unexpected. When prompted, almost half of our respondents (44%) acknowledged that they did not feel they sufficiently understand probabilities and would not be able to properly testify. This was further evident in some of the free-text responses to other questions. For example:

> 'Probabilistic language is currently very confusing to me and I would need a lot more training and understanding of it, before I would be comfortable putting it in a report and testifying to it in a courtroom setting (117).
>
> I agree that the shift toward probabilistic language is appropriate but I still don't fully understand the impact at this time and have had no training on the subject of probabilistic language yet (108).
>
> For the latent examiner not as comfortable with explaining statistics and probability, it could open the door for the attorney to discredit the examiner. That is not only problematic for the case it could be detrimental for the examiner's career (103).'

We are sympathetic to this concern because it has never been a formal requirement for practitioners to have any background knowledge in principles of probabilities and statistics. This presents yet another challenge: if practitioners are expected to testify using probabilistic language, it will require a coordinated investment by forensic science administrators, educators and trainers to ensure practitioners have the fundamental education and training on probabilistic and statistical principles so that they understand what they are reporting and feel comfortable and confident in their own knowledge on the subject. For example, as one respondent suggested:

> 'As a scientist, I understand why the community cannot testify to absolutes in terms of "identification". The reasoning behind that argument is sound. However, I think it will take a change in the dogma of the science to get practitioners to 1) understand probabilities and what they're actually reporting (and that their ultimate conclusion is not actually changing), and 2) want to change how they testify/report their findings (because they are accustomed to "this is how I've always done it") (30).'

It will also require outreach to attorneys and judges so that they understand the transition and what it means. This will require more than a mere policy-change; it will require a commitment by forensic science administrators, educators, trainers, practitioners and policy-makers to address the foundational gaps in education and training curricula as well as establish operational environments that are conducive to allowing practitioners to explore what it means to report probabilistically and how to do so in a way that they are comfortable with. This is important because the transition to probabilistic reporting should not be done in haste. Fortunately, though, probabilistic reporting, as we define it, does not necessarily require the use of numerical quantities, algorithms and other statistical tools or the reporting of evidence along a full continuum, although those measures are preferred by some. Instead, it can be achieved by simply avoiding claims that an individual is 'the' source of an impression or using terms that imply certainty for single-source attribution. As

practitioners and stakeholders gain comfort with reporting using probabilistic language for comparisons which would normally be categorized as 'identification' under traditional categoric reporting schemes, it can be expanded along the continuum to include more marginal comparisons that still provide useful information, but do not warrant stronger conclusions. The use of numerical quantities, algorithms, and other statistical tools will provide more precise information related to the strength of the evidence, but this transition need not be done in a single act nor contingent upon the availability of such technologies.

Funding

This research was funded in part by the Center for Statistical Applications in Forensic Evidence sponsored by the National Institute of Standards and Technology [70NANB15H176].

Acknowledgements

The authors wish to express a sincere appreciation to the survey participants. Their candid responses and willingness to provide such rich feedback is critical for forensic statisticians, forensic practitioners, laboratory administrators, legal scholars, social scientists and others interested in improving education and implementation efforts when promoting the adoption of probabilistic reporting standards. The authors are also grateful to Glenn Langenburg, Eric Ray, Michele Triplett, Sandra Siegel and the members of the Friction Ridge Subcommittee of the NIST Organization of Scientific Area Committees for Forensic Science for facilitating distribution of the survey. For research assistance, we are grateful to Matt Barno. The authors are also appreciative of Christophe Champod and Brandon Garrett for their review and editorial feedback on this manuscript. A draft of this article was presented to the University of California, Irvine section of CSAFE. For helpful comments, we are grateful to Hal Stern and the other members of the group.

Appendix A

Dear Analyst, I am an investigator for CSAFE, the Centre for Statistical Applications in Forensic Evidence (http://forensicstats.org/), and a Professor in the Department of Criminology, Law & Society at the University of California, Irvine. CSAFE has partnerships with many forensic laboratories including the Midwest Forensic Resource Centre. As you probably know, CASFE was named a Forensic Science Centre for Excellence by the National Institute of Standards and Technology to 'support NIST's efforts to advance the utility of probabilistic methods to enhance forensic analysis'. As part of this mission, we are conducting a survey of forensic practitioners about current practices regarding probabilistic reporting of forensic results. We are particularly interested in your experiences with probabilistic reporting. We request your participation in this survey. You are eligible to participate in this study if you are 18 years of age or older and a forensic practitioner. Please note that your participation in this study is completely voluntary and declining to participate will not in any way affect your standing with your employer. There are no direct benefits from participation in the study. However, this study may help us better understand the impact of probabilistic reporting on practitioners. We expect the survey to take around 60 minutes. We thank you in advance for your cooperation.

Appendix B

The closed-response survey question in which participants were asked to choose which of three sample statements most closely resembled the wording they currently used in reports of an association between two friction ridge impressions.

Question: Which of the following most closely resembles the language used in your written examination reports when reporting an association between two prints in your practice?

A. *'The latent print on Exhibit ## was identified to the standards bearing the name XXXX".// **OR**//'There is sufficient quality and quantity of detail in agreement to conclude that the latent print on Exhibit ## and the standards bearing the name XXXX originated from the same source. '//**OR**//"The latent print on Exhibit ## matched the standards bearing the name XXXX'.*

B. *'The latent print on Exhibit ## and the standards bearing the name XXXX have corresponding ridge detail. The likelihood of observing this amount of correspondence when two impressions are made by different sources is considered extremely low.'//**OR**// 'The latent print on Exhibit ## and the standards bearing the name XXXX have corresponding ridge detail. This amount of correspondence provides extremely strong support for the proposition the two impressions were made by the same source rather than by different sources.'//**OR**// 'The latent print on Exhibit ## and the standards bearing the name XXXX have corresponding ridge detail. The probability of observing this amount of correspondence is approximately ## times greater when impressions are made by the same source rather than by different sources.'*

C. *'The features within the impressions are clear and there is an abundant amount of consistency (much more than needed to arrive at a conclusion) with no inconsistencies. The conclusion is easily demonstrable to others and others would be hard pressed to find any reasons to doubt the conclusion. There is a large amount of data that infers that the latent impression was deposited by XXXX.'*

Appendix C

A full listing of all the statements for which researcher coding conflicted with self-report.

Number	Statements self-reported as 'categorical' but coded by researchers as 'probabilistic'
1.	*Item ###. The partial latent print of value has characteristics in agreement with the fingerprint/palm print impressions of XXX, Item ###, (finger # & label/left or right palm). In the opinion of this examiner the likelihood that the impressions were made by a different source other than the one listed is very small.*
2.	*I am of the opinion that the latent prints from item #X and exemplar fingerprints from XXX likely originated from the same source.*

Number	Statements self-reported as 'probabilistic' but coded by researchers as 'categorical'
1.	*Amount of agreement between two impressions is compelling, inferring both originated from the same individual (i.e. it is not plausible that the impression originated from a different source).*
2.	*'The likelihood the ff. print impressions were made by another source is so remote that it is considered a practical impossibility'*
3.	*Print Quality: Sufficient for a comparison and an ABIS search. The XX finger of (name/DOB/record#) was identified as having made this print. This is followed by a qualifying statement: Identification: the decision by a qualified examiner, that two friction ridge impressions originated from the same source. The features present in the two impressions are in sufficient correspondence, and the probability the questioned impression was made by a different source is so small, it is negligible.*
4.	*he latent print on exhibit ## and the exemplars bearing the name XXX have corresponding ridge detail. The likelihood of encountering this much correspondence between two different individuals is considered extremely low and I have discounted it.*
5.	*The likelihood the following print impressions were made by another (different) source is so remote that it is considered as a practical impossibility:*
6.	*The impression on the latent card was made by the same source as the known fingerprint/palm print standards as XXXX*
7.	*Prints recovered from Item XXXX have been analysed, compared evaluated and identified with the known prints of XXXX. The card bearing the know prints of XXXX used in this identification was obtained on 'DATE' by 'NAME', an employee of 'WORKPLACE'. Source identification is the opinion of the examiner that two friction ridge skin impressions originated from the same source. This opinion is the decision that the features are in sufficient correspondence and that the probability the questioned impression was made by a different source is so small that it is negligible.*
8.	*The latent lift labelled L-1 was labelled as being lifted from the _____ by Officer ____. It is my conclusion that the latent lift labelled L-1 was made by the right thumb of _____.*

(*Continued*)

9.	*Identification to Name, SID # along with this statement, An identification does not necessarily eliminate the possibility that another person in the world could leave a print with areas of similar agreement. Identification means that within the examiner's experience and knowledge, no other prints with this much similarity have come from different people.*
10.	*The latent prints of comparison value were compared to the standards for XXX with the following conclusions: F1A—Identified to the right middle finger*
11.	*Latent print card 1 of 1, collected from 'insert', has been identified to the right thumb of the fingerprint cared bearing the name XYZ*
12.	*The latent print, L-X, was identified as the XXXX XXXXX of XXXX XXXX XXXX.*
13.	*Latent impression on lift 0450 originated from the same source as the known right index finger of . . .. . .*
14.	*I formed the opinion that fingerprint R1 and the known fingerprint of XXXX were made by the same finger.*
15.	*Latent # is identified as the number . . . finger of John Doe*
16.	*A visual examination of LL#1 and the ### finger of fingerprint card bearing the name XXXX have sufficient detail to conclude an identification.*
17.	*"identified"*
18.	*A match was made between John Doe ********* and the latent print lifted from the inside passenger door's window by **********. The match was verified by **********
19.	*Suitable detail within 'R2', has been compared with the fingerprint images associated with a Form RCMP C- 216/booking record dated 2017-04-27 bearing the name and particulars of subject. . . SUBJECT KO ##### d.o.b. 1992-05-21 . . .specifically with the impression image(s) of the LEFT THUMB. In the Evaluation phase of the ACE-V process, I considered all of the information gathered during Analysis and Comparison to reach conclusions about the origin of the latent print. As a result of the evaluation of corresponding ridge features observed during comparison, the writer concluded INDIVIDUALIZATION (2), in other words that the impressions from both sources (R2 <> Lt. Thumb print from record of SUBJECT) were caused by the same donor person. The Verification step of the ACE-V (1) process consists of an independent and blind application of the ACE process by a subsequent examiner to either support or refute the conclusions of the original examiner. In the case of the individualized impression(s) above, a vetted version of the case file was forwarded to the FSU Blind Verification Coordinator for assignment to a second analyst for verification or invalidation. D/Cst. *******conducted this review and reported conclusion(s) back to the coordinator. There were no conflicts of opinion. *Note #1***"Notice(s) of Intention—Expert Opinion/Report" under s.657.3 CCC should be served upon the accused or his/her counsel at least 30 days prior to trial date for the introduction of expert testimony. See images for Draft copy of this Form. ***Note #2***Should the introduction of fingerprint evidence be required at trial, the following individuals should be included on the list of witnesses; Cst. ******** - recovered prints—submitted for analysis S/Cst. ********* - analysed, compared, identified fingerprint R2 D/Cst *****conducted independent comparison process verifying conclusion(s) S/Cst. tba- Cell Officer who fingerprints the subject,*

(*Continued*)

*if arrested, for this general occurrence.***Note #3***The presence of a friction ridge print on an item of evidence indicates contact was made between the source and the item of evidence. The presence of a friction ridge print alone does not necessarily indicate the significance of either the contact or the time frame during which the contact occurred.(1) ACE-V—The acronym for a scientific method: Analysis, Comparison, Evaluation, and Verification (see individual terms at* www.swgfast.org*).(2) Individualization -The decision by an examiner that there are sufficient features in agreement to conclude that two areas of friction ridge impressions originated from the same source (donor). Individualization of an impression to one source is the decision that the likelihood the impression was made by another (different) source is so remote that it is considered as a practical impossibility.*

20. *The writer has compared fingerprint impression"R#" described as from 'location of impression', with the fingerprint images appearing on a form RCMP-C216 bearing the name, particulars, and portrait image of Name (DOB: YYYY-MM-DD) KO #######, FPS ########. As a result of the evaluation of corresponding detail observed during this comparison phase, the writer has concluded individualization' in other words that the fingerprints from the aforementioned sources, the latent print 'R#', and the 'specified digit' fingerprint purported to be that of Name, were caused by the same donor-person. Individualization—The decision by an examiner that there are sufficient features in agreement to conclude that two areas of friction ridge impressions originated from the same source (donor). Individualization of an impression to one source is the decision that the likelihood the impression was made by another (different) source is so remote that it is considered as a practical impossibility.*
21. *the Latent print was identified to XXXX*
22. *impression x was identified as the such and such finger/palm of John Doe (sid#). dentification is the opinion of an examiner that there is sufficient quality and quantity of detail in agreement to conclude that two impressions originated from the same source.*
23. *Latent print X was identified to the exemplars bearing the name X*
24. *I compared the latent print from card (#) to the tenprint card of (subjects name) using the ACE-V method. I identified the latent print as being made by the (# finger/palm) of (subjects name). The identification was verified by (Forensic Specialist).*
25. *Latent(s) XX was/were compared to the exemplar prints bearing the name(s) of the above listed subject(s) and was/were IDENTIFIED in the following manner:*
26. *Comparison of the latent print to the known prints listed above revealed there is sufficient information in agreement based upon features, sequence, and spatial relationship to conclude that the latent fingerprint of Lab Item #1 was made by the same individual whose known prints appear on Lab Item 2.*

References

AAAS (2017). Forensic Science Assessments: A Quality and Gap Analysis–Latent Fingerprint Examination. American Association for the Advancement of Science.

Aitken, C., Roberts, P. and Jackson, G. (2010). Fundamentals of Probability and Statistical Evidence in Criminal Proceedings: Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses. Royal Statistical Society.

Ashbaugh, D. R. (1999). *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*. CRC Press: Boca Raton, Fla.

Bali, A. S., Edmond, G., Ballantyne, K. N., Kemp, R. I. and Martire, K. A. (2020). Communicating Forensic Science Opinion: An Examination of Expert Reporting Practices. *Science & Justice* **60**, 216–224.

Bayer, D., Neumann, C. and Ranadive, A. (2016). Communication of Statistically Based Conclusions to Jurors—A Pilot Study. *Journal of Forensic Identification* **66**(5), 405–427.

Broeders, A. P. A. (2006). Of Earprints, Fingerprints, Scent Dogs, Cot Deaths and Cognitive Contamination–A Brief Look at the Present State of Play in the Forensic Arena. *Forensic Science International* **159**, 148–157.

Bush, L. (2009). The Authority of Fingerprint Experts: Is it Based on Belief or Science? *Journal of Forensic Identification* **59**(6), 599–608.

Campbell, A. (2011). The Fingerprint Inquiry Report. The Scottish Government. www.thefingerprintinquiry scotland.org.uk (accessed 13 February 2021).

Champod, C. and Evett, I. W. (2001). A Probabilistic Approach to Fingerprint Evidence. *Journal of Forensic Identification* **51**(2), 101–122.

Chugh, T., Cao, K., Zhou, J., Tabassi, E. and Jain, A. K. (2018). Latent Fingerprint Value Prediction: Crowd-based Learning. *IEEE Transactions on Information Forensics and Security* **13**(1), 20–34.

Cole, S. A. (2001). *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Harvard University Press: Cambridge.

Cole, S. A. (2007). Where the Rubber Meets the Road: Thinking About Expert Evidence as Expert Testimony. *Villanova Law Review* **52**, 803–842.

Cole, S. A. (2009). Forensics without Uniqueness, Conclusions without Individualization: The New Epistemology of Forensic Identification. *Law, Probability and Risk* **8**(3), 233–255.

Cole, S. A. (2014). Indvidiualization Is Dead, Long Live Individualization! Reforms of Reporting Practices for Fingerprint Analysis in the United States. *Law, Probability and Risk* **13**(2), 117–150.

Cole, S. A. (2018). A Discouraging Omen: A Critical Evaluation of the Approved Uniform Language for Testimony and Reports for the Forensic Latent Print Discipline. *Georgia State University Law Review* **34**(4), 1103–1128.

CSAFE (n.d.). Forensic Science Data Portal. Center for Statistical Applications in Forensic Evidence. https:// forensicstats.org/data/ (accessed 13 February 2021).

Defense Forensic Science Center (2015). Use of the term 'Identification' in Latent Print Technical Reports. United States Department of the Army. http://onin.com/fp/DFSC_LP_Information_Paper_Nov_ 2015.pdf (accessed 13 February 2021).

Defense Forensic Science Center (2017). Modification of Latent Print Technical Reports to Include Statistical Calculations. United States Department of the Army. https://osf.io/8kajs/download (accessed 13 February 2021).

Egli, N., Anthonioz, A. and Champod, C. (2014). Evidence Evaluation in Fingerprint Comparison and Automated Fingerprint Identification Systems: Modeling Between Finger Variability. *Forensic Science International* **235**, 86–101.

Egli, N., Champod, C. and Margot, P. (2007). Evidence Evaluation in Fingerprint Comparison and Automated Fingerprint Identification Systems – Modeling with Finger Variability. *Forensic Science International* **167**(2–3), 189–195.

ELDRIDGE, H. (2017). The shifting landscape of latent print testimony: an american perspective. *Journal of Forensic Science and Medicine* **3**, 72–81.

ELDRIDGE, H. (2019). Juror Comprehension of Forensic Expert Testimony: A Literature Review and Gap Analysis. *Forensic Science International: Synergy* **1**, 24–34.

ENFSI (2015). Guideline for Evaluative Reporting in Forensic Science. European Network of Forensic Science Institutes. http://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf (accessed 13 February 2021).

EVETT, I. W. (1995). Avoiding the Transposed Conditional. *Science & Justice* **35**(2), 127–131.

EVETT, I. W. (2015). The Logical Foundations of Forensic Science: Toward Reliable Knowledge. *Philosophical Transactions of the Royal Society B* **370**, 1674.

GARRETT, B. L. and MITCHELL, G. (2013). How Jurors Evaluate Fingerprint Evidence: The Relative Importance of Match Langauge, Method Information, and Error Acknowledgement. *Journal of Empirical Legal Studies* **10**(3), 484–511.

GARRETT, B. L., MITCHELL, G. and SCURICH, N. (2018). Comparing Categorical and Probabilistic Fingerprint Evidence. *Journal of Forensic Sciences* **63**(6), 1712–1717.

GARRETT, R. (2009). IAI Letter Re: NAS Report. To P. J. Leahy. Mar. 18. http://www.clpex.com/Legacy/TheDetail/300-399/TheDetail397.htm (accessed 13 February 2021).

HICKLIN, R. A., BUSCAGLIA, J. and ROBERTS, M. A. (2013). Assessing the Clarity of Friction Ridge Impressions. *Forensic Science International* **226**, 106–117.

IAI (1979) Resolution VII. Identification News 29, International Association for Identification, 1.

IAI (1980) Resolution V. Identification News 30, International Association for Identification, 3.

IAI (2010) Resolution 2010-18. International Association for Identification. July 16.

JAYAPRAKASH, P. T. (2013). Practical Relevance of Pattern Uniqueness in Forensic Science. *Forensic Science International* **231**, 403e401–403e416.

KALKA, N., BEACHLER, M. and HICKLIN, R. A. (2020). LQMetric: A Latent Fingerprint Quality Metric for Predicting AFIS Performance and Assessing the Value of Latent Fingerprints. *Journal of Forensic Identification* **70**(4), 443–463.

KELLMAN, P. J., MNOOKIN, J. L., ERLIKHMAN, G., GARRIGAN, P., GHOSE, T., METTLER, E., CHARLTON, D. and DROR, I. E. (2014). Forensic Comparison and Matching of Fingerprints: Using Quantitative Image Measures for Estimating Error Rates through Understanding and Predicting Difficulty. *PLoS ONE* **9**(5), e94617.

KOEHLER, J. J. (2016). Intuitive error rate estimates for the forensic sciences. *Jurimetrics* **57**, 153.

LANGENBURG, G., NEUMANN, C., MEAGHER, S. B., FUNK, C. and AVILA, J. P. (2013). Presenting Probabilities in the Courtroom: A Moot Court Exercise. *Journal of Forensic Identification* **63**(4), 424–488.

LEEGWATER, A. J., MEUWLY, D., SJERPS, M., VERGEER, P. and ALBERINK, I. (2017). Performance Study of a Score-based Likelihood Ratio System for Forensic Fingermark Comparison. *Journal of Forensic Sciences* **62**(3), 626–640.

LIEBERMAN, J. D., CARRELL, C. A., MIETHE, T. D. and KRAUSS, D. A. (2008). Gold versus Platinum: Do Jurors Recognize the Superiority and Limitations of DNA Evidence Compared to Other Types of Forensic Evidence? *Psychology, Public Policy, and Law* **14**(1), 27–62.

MCKASSON, S. (2001). I Think Therefore I *Probably* Am. *Journal of Forensic Identification* **51**(3), 217–221.

MEUWLY, D. (2006). Forensic Individualisation from Biometric Data. *Science & Justice* **46**(4), 205–213.

MORRISON, G. S., SAHITO, F. H., JARDINE, G., DJOKIC, D., CLAVET, S., BERGHS, S. and DORNY, C. G. (2016). INTERPOL Survey of the Use of Speaker Identification by Law Enforcement Agencies. *Forensic Science International* **263**, 92–100.

NEUMANN, C., CHAMPOD, C., PUCH-SOLIS, R., EGLI, N., ANTHONIOZ, A. and BROMAGE-GRIFFITHS, A. (2007). Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae. *Journal of Forensic Sciences* **52**(1), 54–64.

Neumann, C., Champod, C., Puch-Solis, R., Egli, N., Anthonioz, A., Meuwly, D. and Bromage-Griffiths, A. (2006). Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Three Minutiae. *Journal of Forensic Sciences* **51**(6), 1–12.

Neumann, C., Champod, C., Yoo, M., Genessay, T. and Langenburg, G. (2015). Quantifying the Weight of Fingerprint Evidence through the Spatial Relationshop, Directions and Types of Minutiae Observed on Fingermarks. *Forensic Science International* **248**, 154–171.

Neumann, C., Evett, I. W. and Skerrett, J. (2012). Quantifying the Weight of Evidence from a Forensic Fingerprint Comparison: A New Paradigm. *Journal of the Royal Statistical Society A* **175**(2), 371–415.

Nill, N. (2007). *IQF (Image Quality of Fingerprints) Software Application*. MITRE.

NIST (2012). Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach. Expert Working Group on Human Factors in Latent Print Analysis, U.S. Department of Commerce. National Institute of Standards and Technology. http://www.nist.gov/customcf/get_pdf.cfm?pub_id=910745 (accessed 13 February 2021).

NRC (2009). Strengthening Forensic Science in the United States: A Path Forward. National Research Council. Committee on Identifying the Needs of the Forensic Science Community. https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf (accessed 13 February 2021).

OSAC FRS (2016). Comment on FR Doc # N/A. To July 8. https://www.regulations.gov/comment/DOJ-OLP-2016-0012-0067 (accessed 13 February 2021).

OSAC FRS (2018). Standard for Friction Ridge Examination Conclusions. Friction Ridge Subcommittee of the Organization of Scientific Area Committees. https://www.nist.gov/system/files/documents/2020/03/23/OSAC%20FRS%20CONCLUSIONS%20Document%20Template%202020_Final.pdf (accessed 13 February 2021).

Page, M., Taylor, J. and Blenkin, M. (2011). Uniqueness in the Forensic Identification Sciences – Fact or Fiction? *Forensic Science International* **206**, 12–18.

PCAST (2016). Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. 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 (accessed 13 February 2021).

Pulsifer, D., Muhlberger, S., Williams, S., Shaler, R. and Lakhtakia, A. (2013). An Objective Fingerprint Quality-grading System. *Forensic Science International* **231**, 204–207.

Ribeiro, G., Tangen, J. M. and McKimmie, B. M. (2019). Beliefs about Error Rates and Human Judgment in Forensic Science. *Forensic Science International* **297**, 138–147.

Robertson, B. W. N. (1990). Fingerprints, Relevance and Admissibility. *New Zealand Recent Law Review* **2**, 252–258.

Saks, M. J. and Koehler, J. J. (2008). The Individualization Fallacy in Forensic Science Evidence. *Vanderbilt Law Review* **61**, 199–219.

Stoney, D. A. (1991). What Made Us Ever Think We Could Individualize Using Statistics? *Journal of the Forensic Science Society* **31**(2), 197–199.

SWGFAST (2013). Standards for Examining Friction Ridge Impressions and Resulting Conclusions Draft for Comment. Scientific Working Group on Friction Ridge Analysis Study and Technology. Apr. 27. http://clpex.com/swgfast/documents/examinations-conclusions/130427_Examinations-Conclusions_2.1.pdf (accessed 13 February 2021).

Swofford, H. J., Champod, C., Koertner, A. J., Eldridge, H. and Salyards, M. J. (2021). A Method for Measuring the Quality of Friction Skin Impression Evidence: Method Development and Validation. *Forensic Science International* **320**, 11073.

Swofford, H. J., Koertner, A. J., Zemp, F., Ausdemore, M., Liu, A. and Salyards, M. J. (2018). A Method for the Statistical Interpretation of Friction Ridge Skin Impression Evidence: Method Development and Validation. *Forensic Science International* **287**, 113–126.

THOMPSON, W. C., GRADY, R. H., LAI, E. and STERN, H. S. (2018). Perceived strength of forensic scientists' reporting statements about source conclusions. *Law, Probability and Risk* **17**(2), 133–155.
THOMPSON, W. C. and NEWMAN, E. J. (2015). Lay Understanding of Forensic Statistics: Evaluation of Random Match Probabilities, Likelihood Ratios, and Verbal Equivalents. *Law & Human Behavior* **39**(4), 332–349.
TRIPLETT, M. (2016). Complexity, Level of Association and Strength of Fingerprint Conclusions. *Journal of Cold Case Review* **1**(2), 6–15.
TRIPLETT, M. (2018). Fingerprint Examination: A Defined Method. *Michele Triplett's Fingerprint Information*, edition 3. http://fprints.nwlean.net/Fingerprint%20Examinations%20-%20A%20Defined%20Method%20-%20version%203.pdf (accessed 13 February 2021).
VAN STRAALEN, E. K., DE POOT, C. J., MALSCH, M. and ELFFERS, H. (2020). The Interpretation of Forensic Conclusions by Criminal Justice Professionals: The Same Evidence Interpreted Differently. *Forensic Science International* **313**, 110331.
WELLS, G. L. (1992). Naked Statistical Evidence of Liability: Is Statistical Probability Enough? *Journal of Personality and Social Psychology* **62**(5), 739–752.
YOON, S., LIU, E. and JAIN, A. K. (2015). On Latent Fingerprint Image Quality. In U. GARAIN and F. SHAFAIT (eds), *Revised Selected Papers*. Proceedings of the 5th International Workshop on Computational Forensics, Tsukuba, Japan, 2012. Springer.
ZHENG, X. A. (2016). Ballistics Toolmark Database. National Institute of Standards and Technology. https://www.nist.gov/programs-projects/nist-ballistics-toolmark-database (accessed 13 February 2021).