

2-2021

Latent print comparison and examiner conclusions: A field analysis of case processing in one crime laboratory

Brett O. Gardner
University of Virginia

Sharon Kelley
University of Virginia

Maddisen Neuman
Houston Forensic Science Center and CSAFE

Follow this and additional works at: https://lib.dr.iastate.edu/csafa_pubs



Part of the [Forensic Science and Technology Commons](#)

Recommended Citation

Gardner, Brett O.; Kelley, Sharon; and Neuman, Maddisen, "Latent print comparison and examiner conclusions: A field analysis of case processing in one crime laboratory" (2021). *CSAFE Publications*. 78. https://lib.dr.iastate.edu/csafa_pubs/78

This Article is brought to you for free and open access by the Center for Statistics and Applications in Forensic Evidence at Iowa State University Digital Repository. It has been accepted for inclusion in CSAFE Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Latent print comparison and examiner conclusions: A field analysis of case processing in one crime laboratory

Abstract

Scholarship on the latent print comparison process has expanded in recent years, responsive to the call for rigorous research by scholarly groups (e.g., National [Academy of Sciences](#), 2009; President's Council of Advisors on Science and Technology, 2016). Important to the task of ultimately improving accuracy, consistency, and efficiency in the field is understanding different workflows and case outcomes. The current study describes the casework completed by a latent print unit in a large laboratory during one calendar year (2018), including a unique workflow that involves *Preliminary AFIS Associations* reported out as investigative leads. Approximately 45% of all examined prints were deemed to be of sufficient quality to enter into AFIS, and 22% of AFIS entries resulted in potential identifications. But examiner conclusions and AFIS outcomes (across three AFIS databases) varied according to case details, print source, and AFIS database. Moreover, examiners differed in case processing, sufficiency determinations, and AFIS conclusions. Results are discussed with respect to implications for future research (e.g., comparing these data to case processing data for other laboratories) and ultimately improving the practice of latent print examination.

Keywords

latent print comparison, case processing, field study, AFIS, examiner differences

Disciplines

Forensic Science and Technology

Comments

The following is a manuscript of an article published as Gardner, Brett O., Sharon Kelley, and Maddisen Neuman. "Latent print comparison and examiner conclusions: A field analysis of case processing in one crime laboratory." *Forensic Science International* 319 (2021): 110642. Posted with permission of CSAFE.

Latent print comparison and examiner conclusions: A field analysis of case processing in one crime laboratory

Brett O. Gardner^{a,*}, Sharon Kelley^a, Maddisen Neuman^{b,c}

^aInstitute of Law, Psychiatry, and Public Policy, University of Virginia, United States

^bHouston Forensic Science Center, United States

^cCenter for Statistics and Applications in Forensic Evidence (CSAFE), United States

ARTICLE INFO

Article history:

Received 9 October 2020

Received in revised form 17 November 2020

Accepted 1 December 2020

Available online 2 December 2020

Keywords:

Latent print comparison

Case processing

Field study

AFIS

Examiner differences

ABSTRACT

Scholarship on the latent print comparison process has expanded in recent years, responsive to the call for rigorous research by scholarly groups (e.g., National Academy of Sciences, 2009; President's Council of Advisors on Science and Technology, 2016). Important to the task of ultimately improving accuracy, consistency, and efficiency in the field is understanding different workflows and case outcomes. The current study describes the casework completed by a latent print unit in a large laboratory during one calendar year (2018), including a unique workflow that involves *Preliminary AFIS Associations* reported out as investigative leads. Approximately 45% of all examined prints were deemed to be of sufficient quality to enter into AFIS, and 22% of AFIS entries resulted in potential identifications. But examiner conclusions and AFIS outcomes (across three AFIS databases) varied according to case details, print source, and AFIS database. Moreover, examiners differed in case processing, sufficiency determinations, and AFIS conclusions. Results are discussed with respect to implications for future research (e.g., comparing these data to case processing data for other laboratories) and ultimately improving the practice of latent print examination.

1. Introduction

Research examining the efficacy and reliability of latent print comparison has expanded in recent years in response to scholars highlighting the absence of empirical support for many forensic science disciplines and calling attention to potential contextual effects in analytic conclusions (e.g., [1,2]). A quite small body of research has attempted to elucidate the error rates of latent print comparison as a forensic discipline. One study in 2011 estimated false positive error rates in latent print comparison conclusions to be approximately 1 in 604 cases [3]. A later, unpublished study discovered a much higher false positive error rate of 1 in 24 cases ([4]; but see Wilkinson et al. [5] for discussion of the accuracy of this estimate). These two studies suggest differing error rates, but no other adequately designed studies have explored the error rate of latent print comparison to date. Indeed, scholars have noted that “conclusions about foundational validity . . . must rest on these two recent studies” ([2], p. 91). Despite the limited research body, the PCAST report ultimately concluded that latent print

comparison is a subjective, but foundationally valid, methodology. However, the report also highlighted the potential for variability among examiner conclusions due to the subjective nature of latent print comparison.

While the field awaits additional studies documenting the error rate of latent print comparison, a growing body of research has examined the influence of contextual effects. This literature suggests that a number of common, task-irrelevant factors can influence conclusions. For example, studies have demonstrated that exposure to highly emotional material [6], suspect confessions [7], other analyst conclusions [8], comparison prints [9], and automated ranking systems [10] can influence examiner decisions, although the extent of this influence varies and is generally more pronounced in ambiguous cases. Latent print comparison relies extensively on human perception and judgment, and repeated findings suggesting that contextual information influences forensic conclusions therefore makes intuitive sense. This is especially concerning in light of recent research indicating that forensic analysts often rely on different types of information when completing cases [11].

In spite of mounting research highlighting the potential for cognitive biases to affect latent print comparison, some scholars have pointed to weaknesses in the extant literature and call for continued research ([12]; cf. [13]). For example, of the five

* Corresponding author at: Institute of Law, Psychiatry, and Public Policy, University of Virginia, Box 800660, Charlottesville, VA, 22908, United States.
E-mail address: bgardner@virginia.edu (B.O. Gardner).

forementioned studies examining the influence of task-irrelevant information on latent print examiner decisions, only one used a sample size larger than 30. More broadly, the research base examining latent print comparison also lacks field studies. The discipline is slowly educating itself regarding the validity of latent print comparison and its susceptibility to extraneous information. But almost no research has examined actual latent print casework to first determine typical analysis procedures and outcomes. What proportion of reviewed prints are deemed suitable for comparison? How often do prints submitted for comparison result in identifications? Moreover, to what extent do productivity and analytic outcomes vary across examiners? These basic questions remain largely unaddressed by the literature. Basic examination of laboratory case processing is sorely needed, yet lacking.

To our knowledge, there have only been two studies of actual outcomes in latent print comparison. Langenburg et al. [14] examined latent print casework in one laboratory during the 2003/2004 and 2009/2010 calendar years. The authors explored the possibility that access to contextual information or interaction with police officers/prosecutors might influence examiner conclusions but found that latent print identification rates typically remained at approximately 21% or 22%. Rairden and colleagues [15] explored two years of case processing data relating to the analysis, comparison, evaluation, and verification (ACE-V) process within one large crime laboratory. The study examined over 12,000 latent prints submitted for review and found that, while most prints were deemed to be of no comparative value (56%), approximately one in four reviewed prints resulted in identifications (26%). The study focused on processes involving multiple examiners (i.e., verification, consultation, and conflict resolution), and highlighted modal outcomes (e.g., an exclusion changed to identification). Findings revealed the absence of a “seniority effect,” that is, there were no differences in the proportion of cases that proceeded to consultation based on the junior or senior status of the verifying examiner. Finally, results suggested patterns of examiner differences with respect to the proportion of cases that proceeded to consultation and the proportion of conclusions changed during the consultation process. However, the findings did not identify potential associations between examiner conclusions and case details.

1.1. Current study

We believe that a foundational aspect of latent print research involves close examination of actual, real-world casework. Research has focused on estimating error rates and identifying potential areas of concern within latent print conclusions made within artificial, controlled environments. But little is known about the actual practice of this discipline. In an effort to expand the literature, we examined actual outcomes from a functioning latent print unit within a large metropolitan crime laboratory.

By tracking key case processing details across time, we sought to replicate core findings from Rairden and colleagues [15] and:

1. Describe the casework completed by latent comparison examiners in a large laboratory over the course of one calendar year (i.e., 2018).
2. Describe the prevalence of examiner conclusions during one year.
3. Explore whether examiner conclusions vary according to casework variables such as anatomical source of a print, offense type, or AFIS database.
4. Explore the extent to which there are examiner differences in examiner conclusions and case processing.

1.2. Method

The Houston Forensic Science Center (HFSC) is a local government corporation that provides forensic services to the city of Houston and other local agencies. HFSC has been accredited by the ANSI-ASQ National Accreditation Board (ANAB) since 2015. We examined all HFSC latent print comparison cases with published reports dated in 2018. Of note, some reports addressed requests that predated 2018, and some work conducted in 2018 is not described in this dataset because the reports were not completed until 2019. We only examined genuine latent print casework and therefore excluded all reports describing simulated casework that were completed as part of the laboratory's blind quality control program (for more information about this program, see [16]). This data was collected by generating a report in JusticeTrax, HFSC's primary laboratory information management system. The report encompassed requests completed by latent print examiners between 1/1/2018 and 12/31/2018 and included such information as case offense type, examiner, sufficiency determination, anatomical source, AFIS database searched, and subsequent results based upon AFIS output.

In total, 17 latent print examiners (12 women, 5 men) submitted reports in 2018. All examiners were certified by the International Association for Identification, and work experience ranged from 5 to 36 years ($M = 15.2$; $SD = 9.6$).

1.3. Latent print comparison procedures

Standard operating procedures within HFSC differ from many laboratories. Upon receipt of a latent print for comparison, examiners first make a determination about the print's sufficiency to be entered into an automated fingerprint identification system (AFIS). At this stage, examiners may conclude that a print: 1) has no comparative value (and further analysis is therefore precluded), 2) has comparative value, but is of insufficient quality to be entered into AFIS, or 3) has comparative value, and is of sufficient quality to be entered into AFIS.

If a latent print is entered into AFIS, examiners subsequently make one of three conclusions depending on the AFIS outcome and further comparison. Examiners may conclude that there is *No Association*, meaning the latent print does not appear to correspond to any print on the AFIS candidate list. Conversely, examiners may, based upon corresponding characteristics between the latent print and the candidate image, conclude that the latent print may have originated from the same source as the candidate image. This conclusion is referred to as a *Preliminary AFIS Association* (PAA). If desired, stakeholders may request a confirmatory comparison of a PAA conclusion, and the latent print is again fully examined by the primary examiner and verified by a secondary examiner before an “official identification” is declared. HFSC emphasizes that PAAs are not identifications and simply represent investigative leads; identifications never result from PAAs alone. At the same time, confirmatory comparisons of PAAs have consistently resulted in official identifications with only one exception in the history of the laboratory. During the data collection period for the current study, 337 PAA conclusions were confirmed, with all resulting in official identifications. Finally, examiners may conclude that an AFIS search resulted in a *Reverse Hit*. HFSC uses this term to describe AFIS entries that resulted in potential matches with prints stored in the respective AFIS unsolved latent file, a repository for unidentified prints from earlier cases. Thus, examiners may indicate potential identifications by concluding, *Preliminary AFIS Association* or *Reverse Hit*, or may indicate that no potential identification exists by concluding, *No Association*.

2. Results

The HFSC latent print comparison unit examined 2975 cases in 2018, with 3239 related requests for analysis.¹ Of the cases, the large majority related to burglary/theft charges ($n = 2057$; 69.1%). An additional 16.3% ($n = 485$) were related to robbery charges, and 4.4% ($n = 130$) were related to homicide charges. The remaining cases addressed a variety of charges (e.g., assault, criminal mischief, drug-related offenses). Approximately one quarter, or 23.7% ($n = 706$), of cases were person offenses.

The HFSC latent print comparison unit examined at least 20,494 prints in 2018.² This means that, on average, each request contained 6.3 prints and/or print cards. However, this varied across cases as some contained a single print and others included as many as 143 prints.

2.1. Examiner conclusions

2.1.1. Sufficiency determinations

Of the 20,494 prints examined by the HFSC latent print comparison unit in 2018, 44.8% ($n = 9177$) were deemed to be of sufficient quality to enter into AFIS. Few prints (1.7%; $n = 353$) were deemed to have comparative value, but to be of insufficient quality to enter into AFIS. Slightly more than half (53.5%; $n = 10,964$) of all prints were determined to have no comparative value. Additionally, 464 print cards were examined but determined to contain no identifiable prints.³

Of the 9530 prints deemed to have comparative value, most were fingerprints (67.7%; $n = 6451$) or palm prints (27.6%; $n = 2631$). Latent print examiners rarely examined joint prints (1.7%; $n = 163$) or unspecified impressions that they could not readily identify as originating from a finger or palm (1.8%; $n = 176$).⁴

2.1.2. AFIS outcomes

HFSC examiners conducted 11,812 searches in automated fingerprint identification systems (AFIS) during 2018. Laboratory policy instructs examiners to first search prints in a county-wide AFIS, entitled MorphoTrak (this AFIS database has since changed its title to IDEMIA AFIS), maintained by the Harris County Sheriff's Office. Prints relating to nonperson offenses (e.g., theft, burglary) are typically only searched in this county-wide AFIS. Other prints that do not produce an association in MorphoTrak are subsequently searched in a statewide AFIS, entitled NEC, which is housed and maintained by Texas Department of Public Safety and accessed through the Federal Bureau of Investigation's (FBI) Universal Latent Workstation software. Finally, remaining prints are searched in the federal AFIS, entitled Next Generation Identification (NGI), which is maintained by the FBI Criminal Justice Information Services Division. Consistent with laboratory policy, most AFIS searches were conducted at the county-level (65.0%; $n = 7680$). State-level

¹ There are more comparison requests than cases because officers often submit additional prints for comparison as cases progress or additional prints become available as they are processed by the latent print processing unit.

² We cannot provide a definitive total number of prints examined due to ambiguity in the number of prints deemed to be of no value. The current data do not indicate whether examiners concluded that a specific print or an entire print card was of no value. Therefore, our estimate of 20,494 prints likely underestimates the true number of prints examined during 2018.

³ Occasionally, print cards are submitted that do not contain any identifiable prints upon analysis. This can be for several reasons. Officers sometimes submit print cards with no identifiable prints as part of standard procedure or because they mistakenly believed an impression was a fingerprint. Additionally, prints are sometimes not properly lifted onto a card due to inadequate technique or difficult surface characteristics.

⁴ We were missing data documenting the anatomical source (e.g., fingerprints, palm impressions) for 109 prints of the 9,530 deemed to have comparative value.

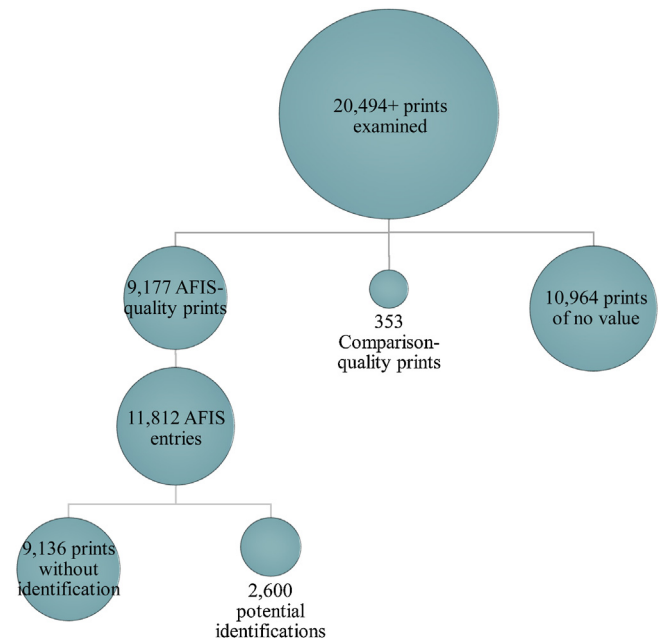


Fig. 1. Flowchart Depicting Prints Examined by the HFSC Latent Print Unit in 2018 and Resulting Outcomes.

Note. There are more AFIS entries that AFIS-quality prints because some prints were searched in multiple AFIS databases.

(i.e. NEC; 16.9%; $n = 1997$) and federal-level (i.e. NGI; 18.1%; $n = 2135$) AFIS searches were equally common to one another.

Of the 11,736 AFIS searches in 2018 for which we have complete data, most did not result in an association (77.8%; $n = 9136$). Indeed, only 20.7% ($n = 2429$) of AFIS entries resulted in what HFSC refers to as a *Preliminary AFIS Association* (PAA). An additional 1.5% ($n = 171$) of AFIS entries resulted in potential matches with prints stored in the unsolved latent file. Again, HFSC refers to these instances as *Reverse Hits*. Fig. 1 provides a flowchart depicting the number of prints processed by HFSC examiners in 2018, and all resulting outcomes in available data. As depicted in the flowchart, approximately 12.7% of all examined prints (i.e., the 20,494 prints examined in 2018) resulted in a potential identification (i.e., PAAs or Reverse Hits).

2.2. Variability within conclusions

2.2.1. Sufficiency determinations

The anatomical source of prints (i.e., fingerprint, palm print, joint print, unspecified impression) was significantly associated with sufficiency determinations among prints deemed to be of comparative value, $\chi^2(3, N = 9421) = 27.67, p < .001$, Cramer's $V = .05$, 95% CI [.03, .08].⁵ Specifically, unspecified impressions were at least 2.5 times more likely to be deemed to be of comparative value but of insufficient quality for AFIS (as opposed to AFIS quality) than were other latent print sources.

Examiners were also slightly more likely to conclude that latent prints were AFIS quality (compared to determinations of *no comparative value* or determinations that prints were of comparative value but of insufficient quality for AFIS entry) in cases involving person offenses than they were in cases of nonperson offenses, $\chi^2(2, N = 20,494) = 27.78, p < .001$, Cramer's $V = .04$, 95% CI [.02, .05]. Specifically, 46.4% of prints from person-offense cases

⁵ 95% confidence intervals for Cramer's V obtained via 1,000 bootstrap replications.

Table 1
Outcomes for Prints Entered into AFIS by HFSC Examiners in 2018.

Anatomical Source	AFIS Outcome			Total
	Preliminary Association	Reverse Hit	No Association	
Fingerprint	22.7% (n = 1411)	2.0% (n = 127)	75.2% (n = 4667)	100% (n = 6205)
Palm print	27.4% (n = 708)	1.5% (n = 39)	71.1% (n = 1837)	100% (n = 2584)
Joint print	4.3% (n = 7)	0.0% (n = 0)	95.7% (n = 156)	100% (n = 163)
Unspecified Impression	4.7% (n = 9)	0.0% (n = 0)	95.3% (n = 182)	100% (n = 191)
Total	23.4% (n = 2135)	1.8% (n = 166)	74.8% (n = 6842)	100% (N = 9143)

Note. N = 9143 prints with available data of the 9177 total prints deemed to be of sufficient quality for AFIS entry.

were determined to be AFIS quality compared to 43.7% of prints from nonperson-offense cases.

2.2.2. AFIS outcomes

Print type was significantly associated with AFIS conclusions, $\chi^2(6, N = 9143) = 107.61, p < .001$, Cramer's $V = .08$, 95% CI [.07, .09]. As Table 1 demonstrates, fingerprints and palm prints were more than 5 times as likely to result in potential identifications as were joint prints or unspecified impressions. Indeed, the vast majority of joint prints and unspecified impressions (~95%) searched in AFIS did not result in potential identifications.

Offense type was also associated with AFIS conclusions. Specifically, AFIS conclusions varied between cases involving person and nonperson offenses, $\chi^2(2, N = 11,736) = 128.85, p < .001$, Cramer's $V = .11$, 95% CI [.09, .12]. Although prints associated with nonperson offenses are generally only searched in the county-wide AFIS, such prints were 1.28 times more likely to result in a potential identification than were prints from person-offense cases (25.0% vs. 19.6%).⁶

We also examined whether AFIS conclusions varied according to the particular AFIS software that was used to compare the print with others. A chi-square analysis indicated that AFIS conclusions varied significantly by software type, $\chi^2(4, N = 11,729) = 891.53, p < .001$, Cramer's $V = .20$, 95% CI [.18, .21]. Specifically, Table 2 indicates that the state-level AFIS, NEC, was least likely to result in a potential identification (5.0%). MorphoTrak and NGI were approximately 5 times more likely to result in a potential identification. Additionally, prints entered into NGI were more than 9 times as likely to result in a Reverse Hit than were prints entered into other AFIS software systems.

2.2.3. Examiner differences

Of the 17 latent print examiners who submitted reports in 2018, three did not complete independent casework at HFSC for longer than one month during 2018 and we therefore limited our analysis of examiners to the remaining 14. Table 3 details differences among HFSC examiners in case processing, sufficiency determinations, and AFIS conclusions. As shown, examiners completed between 12 and 46 requests each month, examining between 66 and 269 prints and/or print cards. Examiners' rates for determining prints to be of sufficient quality for AFIS entry also appeared to vary. Some examiners opined that only one of every three examined prints (35.8%) were of sufficient quality whereas others opined that more than half of all prints (56.5%) were of sufficient quality. Finally, there was some variability among examiners in the rate of PAAs among prints entered into AFIS. Some examiners

concluded PAAs for only 13.3% of entered prints whereas others were approximately two times as likely (27.1%) to conclude that a PAA existed. While it is certainly possible that examiners do not review equitable caseloads (i.e., some examiners are assigned more difficult cases or cases involving lower-quality prints), we have no reason to suspect that discrepant caseloads completely explain this pattern. Thus, we suspect that examiner differences (i.e., different thresholds for determining prints to be of sufficient quality) account for at least some of this variability, although conclusions regarding this explanation remain limited at present.

Although the limited number of examiners preclude formal statistical analyses, Table 3 also suggests a potential relationship between examiner productivity and conclusions. Specifically, examiners who completed more requests every month appeared less likely to determine prints to be of sufficient quality for AFIS entry and more likely to conclude prints to be of no comparative value. Thus, some amount of the variability in examiner efficiency may be associated with individual differences in sufficiency determinations. Indeed, examiners assigned cases involving prints of lower quality that are not entered into AFIS would certainly be expected to complete more requests than examiners assigned cases requiring multiple AFIS searches.

3. Discussion

The results above reflect one year of cases that proceeded through one latent print unit in a large crime laboratory. There are some noteworthy similarities and differences in the workload and workflow in comparison to Rairden and colleagues' [15] findings documenting latent print case processing in the same laboratory during years 2014–2016. For instance, the number of cases in 2018 (2975) was nearly equivalent to two years of cases (2535) in the earlier study. Thus, workload appears to have increased dramatically. In 2016, approximately 2500 latent lift cards for which analysis was never requested were discovered at the Houston Police Department Property Room and transferred to HFSC for analysis. Furthermore, 17 examiners were employed in 2018 compared to 12 examiners from the previous study. The increase in examiners also likely contributed to increased productivity. With respect to type of cases, the overwhelming majority in both studies were from burglary or robbery offenses.

HFSC workflow has undergone significant changes since the Rairden et al. study, most prominently in the use of PAAs, making some workflow comparisons impossible. Interestingly though, there was remarkable consistency in sufficiency determinations: ~45% of prints were deemed to be of sufficient quality to enter into AFIS in both studies. These findings underscore the importance of this threshold determination if approximately one out of two prints is deemed not to be valuable for comparison or entry into an AFIS. Additionally, the fact that examiners were slightly more likely

⁶ Prints associated with nonperson offenses were 0.27 times less likely to result in Reverse Hits than were prints associated with person offenses (0.6% vs. 2.2%).

Table 2
Outcomes for Prints Entered into AFIS Software Types by HFSC Examiners in 2018.

AFIS Software	AFIS Outcome			Total
	Preliminary Association	Reverse Hit	No Association	
MorphoTrak (county-level)	25.5% (n = 1952)	0.2% (n = 18)	74.3% (n = 5686)	100% (n = 7656)
NEC (state-level)	4.3% (n = 84)	0.7% (n = 14)	95.0% (n = 1846)	100% (n = 1944)
NGI (nationwide-level)	18.4% (n = 392)	6.5% (n = 139)	75.1% (n = 1598)	100% (n = 2129)
Total	20.7% (n = 2428)	1.5% (n = 171)	77.8% (n = 9130)	100% (N = 9143)

Note. N = 9143 prints with available data of the 9177 total prints deemed to be of sufficient quality for AFIS entry.

Table 3
Individual Differences in Case Processing, Sufficiency Determinations, and AFIS Conclusions among HFSC Latent Print Examiners.

Examiner	Case Processing				Sufficiency Determination			AFIS Conclusion			
	Months Employed	Requests	Requests/ Month	Prints/cards Examined	Prints/ Month	% AFIS Qual.	% Not AQ	% NLoV	% PAA	% Reverse Hit	% No Hit
A	4.7	118	25.1	747	158.9	37.8%	0.1%	62.1%	22.4%	0.7%	76.9%
B	12	155	12.9	1201	100.1	56.5%	0.3%	43.1%	17.8%	2.3%	79.9%
C	12	336	28.0	1862	155.2	45.0%	3.3%	51.8%	27.1%	0.0%	72.9%
D	12	220	18.3	1209	100.8	48.5%	1.3%	50.2%	17.3%	2.2%	80.5%
E	12	172	14.3	1121	93.4	44.6%	0.2%	55.2%	13.3%	2.8%	83.9%
F	11	254	23.1	1411	128.3	40.7%	1.5%	57.8%	25.4%	2.5%	72.2%
G	12	146	12.2	794	66.2	44.3%	1.0%	54.7%	16.8%	4.2%	79.1%
H	12	206	17.2	1197	99.8	40.7%	2.7%	56.6%	15.2%	0.6%	84.2%
I	12	550	45.8	3222	268.5	38.6%	0.2%	61.1%	24.5%	1.7%	73.8%
J	7	149	21.3	980	140.0	35.8%	0.1%	64.1%	13.7%	0.2%	86.0%
K	12	136	11.3	1134	94.5	52.3%	0.4%	47.4%	23.0%	1.4%	75.6%
L	12	293	24.4	2248	187.3	46.6%	5.4%	48.0%	18.9%	0.4%	80.7%
M	11.3	217	19.2	1576	139.5	50.4%	2.0%	47.7%	20.4%	0.5%	79.1%
N	12	178	14.8	1221	101.8	49.8%	1.7%	48.5%	22.8%	1.5%	75.8%
Total	M = 11		M = 20.6		M = 141.5	44.8%	1.7%	53.5%	20.7%	1.5%	77.8%

Note. We excluded three examiners because they only completed independent casework for a single month during the data collection period (i.e., 2018). % AFIS Qual. = percentage of examined prints determined to be of sufficient quality for AFIS entry. % Not AQ = percentage of examined prints determined to be of comparative value, but insufficient quality for AFIS entry. % NLoV = percentage of examined prints determined to be of no comparative value. % PAA = percentage of Preliminary AFIS Association conclusions among prints entered into AFIS. % Reverse Hit = percentage of Reverse Hit conclusions among prints entered into AFIS. % No Hit = percentage of prints entered into AFIS that were concluded to not correspond to any print on the AFIS candidate list.

to deem prints in person-offense cases to be of sufficient quality for AFIS entry suggests the possibility of a slight implicit bias based on knowledge of case details (i.e., unintentionally setting a lower threshold for AFIS searches in person-offense cases because they are more important or serious). Alternatively, examiners might have developed an informal but explicit practice—short of formal policy—that more questionable prints should nevertheless be entered into an AFIS for person-offense cases.

Results indicate that HFSC most commonly employed county-level AFIS searches (65%), with similar rates of state-level and federal-level AFIS searches (i.e., about 17% and 18%, respectively). Most AFIS searches did not result in an association—only 21% of AFIS entries led to a PAA. Analyses revealed, however, meaningful differences in rates of PAAs depending on which AFIS was used. Specifically, results indicated that the county and federal AFIS databases were five times more likely than the state-level AFIS to result in a potential identification. Perhaps unsurprisingly, examiners had more success entering fingerprints and palm prints in AFIS compared to joint prints or unspecified impressions.

Results also demonstrated a relationship between case type, examiner decisions, and AFIS outcomes. Although examiners were more likely to determine that prints were AFIS quality in person-offense cases rather than nonperson-offense cases, prints from nonperson offenses were more likely to result in a PAA, perhaps due to the higher rate of success associated with the county-level AFIS (used for most nonperson offenses).

Differences across examiners emerged with respect to both sufficiency determinations and rates of PAAs. Of course, some of these differences are almost certainly driven by differences in cases assigned to examiners. However, the magnitude of the differences (e.g., 35.8% vs. 56.5% of prints determined to be of sufficient quality for AFIS entry) suggests meaningful examiner differences in how they approach decisions (e.g., in tendencies toward deeming prints sufficient for AFIS). This possibility is consistent with the subjective nature of latent print identifications and previous research on lack of consensus at this step of the ACE-V process (e.g., [3]). The fact that examiners who completed more requests each month seemed less likely to determine prints to be sufficient for AFIS entry might suggest that more productive examiners have a higher threshold for sufficiency, perhaps based on more experience or expertise on the quality necessary for an AFIS search. Alternatively though, examiners with higher sufficiency thresholds (or examiners assigned cases with lower-quality prints) may process greater numbers of prints simply because of the additional effort and time associated with AFIS entries.

Certainly, the results of any study can only be interpreted in the context of the study's limitations. As these cases and outcomes are from a single unit in a single year, results cannot be generalized to other laboratories, particularly laboratories that differ in key respects from HFSC (e.g., in size, population served, types of cases). Additionally, the new workflow (i.e., the use of PAAs) limits our ability to review sources of examiner disagreement that would

come from more routine use of verification. As reported above, in only one case did a confirmatory comparison of a PAA result in a determination that the prints did not come from the same source.

In the future, comparing the workflow and outcomes of multiple laboratories would offer greater context for these data and enrich the field's understanding about differences in case processing and case outcomes. Indeed, these results are valuable as a point of comparison to ultimately shed light on procedures that can optimize efficiency and accuracy.

These results also speak to the need for research on AFIS databases. A number of factors might account for different rates of potential identifications generated by the different AFIS databases used by HFSC (i.e., local, state, federal). Although some research endeavors are limited due to the proprietary nature of AFIS algorithms, the National Institute of Standards and Technology has conducted comparisons of widely used AFIS algorithms using standardized latent prints and exemplars. It is likely that identification rate differences among AFIS databases are explained, in part, by different algorithms but also by differences in the number and type of prints contained within AFIS-specific databases. In any case, this finding highlights an under-researched area that has the potential to significantly influence case outcomes. Future research should systematically explore differences in the accuracy of common AFIS databases and the reasons for these differences.

In sum, we consider these results crucial to the aim of increasing transparency in, and dissemination of, crime laboratory case processing information. Almost no research has examined typical analysis procedures and outcomes among latent print casework. The current results suggest that approximately half of examined prints are determined to be of no comparative value, and approximately 13% of examined prints result in potential identifications, with noted variability relating to examiner differences, case details, print source, and AFIS database. We hope these findings serve as a launching point for future research on the latent print comparison process with the aim of improving consistency, efficiency, and accuracy in the discipline.

Funding

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

Declaration of Competing Interest

The authors report no declarations of interest.

CRediT authorship contribution statement

Brett O. Gardner: Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Sharon Kelley:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Maddisen Neuman:** Investigation, Data curation, Writing - original draft, Writing - review & editing.

References

- [1] National Academy of Sciences, Strengthening Forensic Science in the United States: A Path Forward, National Academies Press, Washington, DC, 2009.
- [2] President's Council of Advisors on Science and Technology, Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, Executive Office of the President of the United States, Washington, DC, 2016.
- [3] B.T. Ulery, A. Hicklin, J. Buscaglia, M.A. Roberts, Accuracy and reliability of forensic latent fingerprint decisions, *PNAS* 108 (2011) 7733–7738, doi:<http://dx.doi.org/10.1073/pnas.1018707108>.
- [4] I. Pacheco, B. Cerchiai, S. Stoiloff, Miami-Dade Research Study for the Reliability of the ACE-V Process: Accuracy & Precision in Latent Fingerprint Examinations, (2014) . Retrieved from www.ncjrs.gov/pdffiles1/nij/grants/248534.pdf.
- [5] D. Wilkinson, D. Richard, D. Hockey, Expert fingerprint testimony post-PCAST—a Canadian case study, *J. Forensic Identif.* 68 (2018) 299–331.
- [6] I.E. Dror, A.E. Péron, S. Hind, D. Charlton, When emotions get the better of us: the effect of contextual top-down processing on matching fingerprints, *Appl. Cognit. Psychol.* 19 (2005) 799–809.
- [7] I.E. Dror, D. Charlton, Why experts make errors, *J. Forensic Identif.* 56 (4) (2006) 600–616.
- [8] G. Langenburg, C. Champod, P. Wertheim, Testing for potential contextual bias effects during the verification stage of the ACE-V methodology when conducting fingerprint comparisons, *J. Forensic Sci.* 54 (2009) 571–582.
- [9] I.E. Dror, C. Champod, G. Langenburg, D. Charlton, H. Hunt, R. Rosenthal, Cognitive issues in fingerprint analysis: inter- and intra-expert consistency and the effect of a 'target' comparison, *Forensic Sci. Int.* 208 (2011) 10–17.
- [10] I.E. Dror, K. Wertheim, P. Fraser-Mackenzie, J. Walajjys, The impact of human-technology cooperation and distributed cognition in forensic science: biasing effects of AFIS contextual information on human experts*, *J. Forensic Sci.* 57 (2012) 343–352.
- [11] B.O. Gardner, S. Kelley, D.C. Murrrie, I.E. Dror, What do forensic analysts consider relevant to their decision making? *Sci. Justice* 59 (2019) 516–523, doi:<http://dx.doi.org/10.1016/j.scijus.2019.04.005>.
- [12] L.J. Curley, J. Munro, M. Lages, An inconvenient truth: more rigorous and ecologically valid research is needed to properly understand cognitive bias in forensic decisions, *Forensic Sci. Int.: Synergy* 2 (2020) 107–109, doi:<http://dx.doi.org/10.1016/j.fsisyn.2020.1.004>.
- [13] J. Kukucka, People who live in ivory towers shouldn't throw stones: a refutation of Curley et al, *Forensic Sci. Int.: Synergy* 2 (2020) 110–113, doi:<http://dx.doi.org/10.1016/j.fsisyn.2020.03.001>.
- [14] G. Langenburg, F. Bochet, S. Ford, A report of statistics from latent print casework, *Forensic Sci. Policy Manage.: Int. J.* 5 (2014) 15–37, doi:<http://dx.doi.org/10.1080/19409044.2014.929759>.
- [15] A. Rairden, B.L. Garrett, S. Kelley, D. Murrrie, A. Castillo, Resolving latent conflict: what happens when latent print examiners enter the cage? *Forensic Sci. Int.* 289 (2018) 215–222, doi:<http://dx.doi.org/10.1016/j.forsciint.2018.04.040>.
- [16] C. Hundl, M. Neuman, A. Rairden, P. Rearden, P. Stout, Implementation of a blind quality control program in a forensic laboratory, *J. Forensic Sci.* 65 (2019) 814–822.