Short Report

Psychological Science
22(8) 995–997

DOI: 10.1177/0956797611414729
http://pss.sagepub.com

Identifying Fingerprint Expertise

Jason M. Tangen[1,2], Matthew B. Thompson[1,2], and Duncan J. McCarthy[3]

[1]School of Psychology, The University of Queensland; [2]National Information and Communications Technology Australia, Queensland Research Laboratory, St. Lucia, Queensland, Australia; and [3]Forensic Services Branch, Queensland Police Service, Brisbane, Queensland, Australia

Received 3/3/11; Revision accepted 4/23/11

"CSI"-style TV shows give the impression that fingerprint identification is fully automated. In reality, when a fingerprint is found at a crime scene, it is a human examiner who is faced with the task of identifying the person who left the print—a task that falls squarely in the domain of psychology. The difficulty is that no properly controlled experiments have been conducted on fingerprint examiners' accuracy in identifying perpetrators (Loftus & Cole, 2004), even though fingerprints have been used in criminal courts for more than 100 years. Examiners have even claimed to be infallible (Federal Bureau of Investigation, 1984). However, the U.S. National Academy of Sciences has recently condemned these claims as scientifically implausible, reporting that faulty analyses may be contributing to wrongful convictions of innocent people (National Research Council, Committee on Identifying the Needs of the Forensic Science Community, 2009).

Proficiency tests of fingerprint examiners and previous studies of examiners' performance have not adequately addressed the issue of accuracy, and they been heavily criticized for (among other things) failing to include large, counterbalanced samples of targets and distractors for which the ground truth is known (see Cole, 2008, and Vokey, Tangen, & Cole, 2009). Thus, it is not clear what these tests say about the proficiency of fingerprint examiners, if they say anything at all. Researchers at the National Academy of Sciences and elsewhere (e.g., Saks & Koehler, 2005; Spinney, 2010) have argued that there is an urgent need to develop objective measures of accuracy in fingerprint identification. Here we present such data.

Method

Participants

Thirty-seven qualified practicing fingerprint experts from five police organizations (the Australian Federal, New South Wales, Queensland, South Australia, and Victoria Police) participated in the study. In addition, 37 undergraduates from The University of Queensland participated for course credit, providing comparison data on the performance of novices.

Procedure

We presented the 37 qualified fingerprint experts and the 37 novices with pairs of prints displayed side by side on a computer screen, as illustrated in Figure 1. Participants were asked to judge whether the prints in each pair matched, using a confidence rating scale ranging from 1 (*sure different*) to 12 (*sure same*); judgments were reported by moving a scroll bar to the left ("different") or right ("same"). Note that the scale forced a "match" or "no match" decision because ratings of 1 through 6 indicated a match, whereas ratings of 7 through 12 indicated no match. Judgments that the information was "inconclusive," which are often made in practice, were not permitted in this two-alternative forced-choice design, so it was possible to distinguish between accuracy and response bias (Green & Swets, 1966). This task emulates the most forensically relevant aspect of the identification process, namely, the extent to which a print can be accurately matched to its source.

Stimuli

The stimuli consisted of 36 simulated crime-scene prints that were paired with fully rolled prints. Across participants, each simulated print was paired with a fully rolled print from the same individual (match), with a nonmatching but similar exemplar (similar distractor), and with a random nonmatching exemplar (nonsimilar distractor). For each participant, each simulated print was randomly allocated to one of the three trial types, with the constraint that there were 12 prints in each condition.

The simulated prints and their matches were from the Forensic Informatics Biometric Repository,[1] so, unlike genuine crime-scene prints, they had a known true origin (Cole, 2005). Simulated prints were dusted by a research assistant (who was trained by a qualified fingerprint expert), photographed, cropped to 600 × 600 pixels, and isolated in the

Corresponding Author:
Jason M. Tangen, School of Psychology, The University of Queensland, St. Lucia, Queensland 4072, Australia
E-mail: jtangen@psy.uq.edu.au

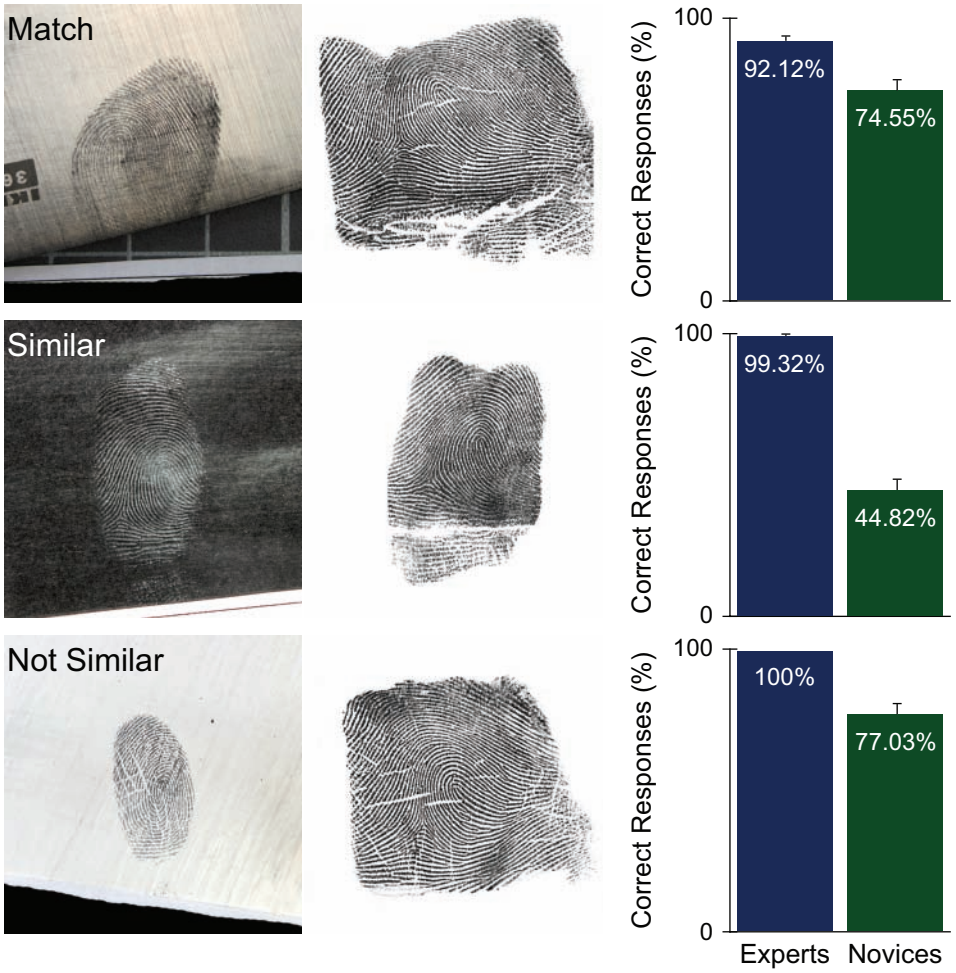

Fig. 1. Stimuli and results. On each trial, participants were presented with a simulated crime-scene print on the left and a fully rolled candidate print on the right, and they were asked to indicate their level of confidence in whether the prints matched. On some trials, the two prints came from the same individual (top row); on others, the prints were similar but came from two different individuals (middle row); and on others, the prints came from two different individuals and were paired randomly (bottom row). The three graphs on the right depict experts' and novices' mean percentage of correct responses in these three conditions. Error bars represent 95% within-cell confidence intervals.

frame. A qualified expert (the third author) reported that each simulated print contained sufficient information to make an identification if there was a clear comparison exemplar. The matching exemplars were fully rolled fingerprint impressions made using a standard elimination pad and a 10-print card. Each card was scanned in color as a 600-dpi lossless Tagged Information File Format (TIFF) file, and each print was cropped to 600 × 600 pixels and isolated in the frame.

Similar distractors were obtained by searching the Australian National Automated Fingerprint Identification System. For each simulated print, the most highly ranked nonmatching exemplar from the search was used if it was available in the Queensland Police 10-print hard-copy archives, which contains approximately 3.3 million prints. The corresponding 10-print card was retrieved from the archives, scanned, and extracted by the same method as before. In practice, highly similar nonmatches retrieved from large national databases are likely to increase the chance of incorrect identifications (Dror & Mnookin, 2010). Distinguishing such highly similar, but nonmatching, prints from genuine matches is potentially the most difficult task that fingerprint examiners face. The nonsimilar distractor for a given simulated print was randomly selected from the entire set of matching and similar distractors after removing the match and similar distractor for that simulated print.

Results

For each participant, we calculated the percentage of trials responded to correctly in each condition. The three graphs on

the right side of Figure 1 depict the average percentage of correct responses for the 37 experts and 37 novices.

As the figure shows, experts performed exceedingly well. On the 12 trials in which the prints matched, experts correctly identified 92.12% of the pairs, on average, as matches (hits), misidentifying 7.88% as nonmatches (misses). Misses are the kind of error that can lead to a failure to identify a criminal.

On the 12 similar-distractor trials, experts correctly declared nearly all of the pairs (99.32%) to be nonmatches (correct rejections); only 3 pairs (0.68%) out of the 444 in this condition were incorrectly declared to be matches (false alarms). Experts did not misidentify any of the 12 nonsimilar distractor prints as matches. Such errors can lead to false convictions.

Even though the novices could reliably distinguish matching and nonmatching prints, they made a large number of errors. In particular, novice participants mistakenly identified 55.18% of the similar, nonmatching distractor prints as matches (the corresponding rate for experts was 0.68%).

We subjected the percentages of correct responses to a 2 (expertise: experts, novices) × 3 (trial type: match, similar distractor, nonsimilar distractor) mixed analysis of variance. The analysis revealed significant main effects of expertise, $F(1, 72) = 416.46$, $MSE = 0.013$, $p < .001$, and trial type, $F(2, 144) = 45.68$, $MSE = 0.011$, $p < .001$, and a significant interaction between the two, $F(2, 144) = 64.32$, $MSE = 0.011$, $p < .001$. Simple-effects analyses revealed a significant benefit of expertise on all trial types—match: $F(1, 72) = 38.49$, $MSE = 0.01$; similar distractor: $F(1, 72) = 476.99$, $MSE = 0.01$; and nonsimilar distractor, $F(1, 72) = 98.46$, $MSE = 0.01$.

Conclusions

We have shown that qualified, court-practicing fingerprint experts are exceedingly accurate compared with novices, but are not infallible. Our experts tended to err on the side of caution by making errors that would free the guilty rather than convict the innocent. Even so, they occasionally made the kind of error that can lead to false convictions. Expertise with fingerprints appears to provide a real performance benefit, but fingerprint experts—like doctors and pilots—make mistakes that can put lives and livelihoods at risk.

Qualified fingerprint examiners now have evidence to legitimately claim specialized knowledge, which may satisfy legal admissibility criteria. It remains unclear, however, how our experiment should affect the testimony of forensic examiners and the assertions that they can reasonably make. The issue is no longer whether fingerprint examiners make errors, but rather how to acknowledge those errors.

We have taken a first step in addressing the call by the National Academy of Sciences for cognitive psychology to establish the limits and levels of performance in forensic science. Considering the central role of humans in forensic identification, the field would benefit from further psychological research. Research on clinical reasoning in medicine, for example, developed over the past 40 years, after it became evident that physicians' decisions too often resulted in adverse consequences for patients. Much has been learned about differences between novice and expert medical practitioners, the influence of cognitive biases in medical decision making, and the most effective ways to incorporate such knowledge into practice. Further research into forensic decision making will help to ensure the integrity of forensics as an investigative tool so that the rule of law is justly applied.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Note

1. A description of the Forensic Informatics Biometric Repository and details about the procedures used to collect the stimuli are available at the Forensic Informatics Biometric Repository Web site (www.fib-r.com).

References

Cole, S. A. (2005). More than zero: Accounting for error in latent fingerprint identification. *Journal of Criminal Law and Criminology*, *95*, 985–1078.

Dror, I. E., & Mnookin, J. L. (2010). The use of technology in human expert domains: Challenges and risks arising from the use of automated fingerprint identification systems in forensics. *Law, Probability & Risk*, *9*, 47–67.

Federal Bureau of Investigation. (1984). *The science of fingerprints: Classification and uses*. Washington, DC: U.S. Government Printing Office.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Loftus, E. F., & Cole, S. A. (2004). Contaminated evidence. *Science*, *304*, 959.

National Research Council, Committee on Identifying the Needs of the Forensic Science Community. (2009). *Strengthening forensic science in the United States: A path forward*. Washington, DC: The National Academies Press.

Saks, M. J., & Koehler, J. J. (2005). The coming paradigm shift in forensic identification science. *Science*, *309*, 892–895.

Spinney, L. (2010). The fine print. *Nature*, *464*, 344–346.

Vokey, J. R., Tangen, J. M., & Cole, S. A. (2009). On the preliminary psychophysics of fingerprint identification. *The Quarterly Journal of Experimental Psychology*, *62*, 1023–1040.