Expertise, memory, and non-analytic cognition in fingerprint matching: Experts can discriminate prints in noise, spaced in time, and in the blink of an eye

Matthew Thompson
Jason Tangen
Duncan McCarthy

The University of Queensland
Queensland Police Service
National ICT Australia
UCLA
Expertise, memory, and non-analytic cognition in fingerprint matching: Experts can discriminate prints in noise, spaced in time, and in the blink of an eye

Matthew Thompson
Jason Tangen
Duncan McCarthy

The University of Queensland
Queensland Police Service
National ICT Australia
UCLA
Development of Expertise
Development of Expertise

Noise and Inversion
Development of Expertise

Noise and Inversion

Memory
Development of Expertise

Noise and Inversion

Memory

Deadline
Development of Expertise
Identifying Fingerprint Expertise

Jason M. Tangen, Matthew B. Thompson, and Duncan J. McCarthy

Psychological Science


“CSF”-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 have been critically 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, 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 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
37 Novices (no experience)
37 Novices (no experience)

9 New Trainees (5 weeks or 6 months)
37 Novices (no experience)

9 New Trainees (5 weeks or 6 months)

8 Intermediate Trainees (1-5 years)
37 Novices (no experience)

9 New Trainees (5 weeks or 6 months)

8 Intermediate Trainees (1-5 years)

37 Experts (5-32 years; $M = 17$)
45 Pairs

Targets (matches)
Casework Materials

Targets
<table>
<thead>
<tr>
<th>Experts</th>
<th>Novices</th>
<th>New Trainees</th>
<th>Intermediate Trainees</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Match</strong></td>
<td><strong>Correct Responses (%)</strong></td>
<td><strong>Match</strong></td>
<td><strong>Correct Responses (%)</strong></td>
</tr>
<tr>
<td>Match</td>
<td>75.32%</td>
<td>99.21%</td>
<td>100%</td>
</tr>
<tr>
<td>Not Similar</td>
<td>43.27%</td>
<td>73.17%</td>
<td>97.24%</td>
</tr>
<tr>
<td>High similarity distractors</td>
<td>69.36%</td>
<td>49.15%</td>
<td>69.38%</td>
</tr>
<tr>
<td>Random distractors (non-matches)</td>
<td>95.42%</td>
<td>73.11%</td>
<td>97.24%</td>
</tr>
</tbody>
</table>

45 Pairs

Targets (matches)
45 Pairs

**Targets (matches)**

- Match 1: 75.32%
- Match 2: 99.21%
- Match 3: 100%

**Not Similar**

- Match 1: 43.27%
- Match 2: 73.17%
- Match 3: 97.24%

**Similar**

- Match 1: 75.32%
- Match 2: 99.21%
- Match 3: 100%

**Highly similar distractors (non-matches)**

- Similar 1: 43.27%
- Similar 2: 73.17%
- Similar 3: 97.24%

**New Trainees**

- New Trainees 1: 69.36%
- New Trainees 2: 49.15%
- New Trainees 3: 69.38%

**Intermediate Trainees**

- Intermediate Trainees 1: 69.36%
- Intermediate Trainees 2: 49.15%
- Intermediate Trainees 3: 69.38%
45 Pairs

Targets (matches)

Highly similar distractors (non-matches)
45 Pairs

Targets (matches)

Highly similar distractors (non-matches)

Random distractors (non-matches)
<table>
<thead>
<tr>
<th>Match</th>
<th>No Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

**Liberal Bias Maximum**

**Conservative Bias Maximum**

**More Liberal**

**More Conservative**

**Less Discriminating**

**More Discriminating**
More Liberal

More Discriminating

Less Discriminating

More Conservative

Liberal Bias Maximum

Conservative Bias Maximum

Initial training increases conservatism, not discriminability
Initial training increases conservatism, not discriminability

Lots of experience doesn’t add much
Initial training increases conservatism, not discriminability

Lots of experience doesn’t add much

Expertise lies in discriminating similar, non-matches
What is the nature of fingerprint expertise?
Norman et al (1989)
Evans & Wolfe, 2012
Evans & Wolfe, 2012
Development of Expertise
Development of Expertise

Noise and Inversion

Memory

Deadline
Noise and Inversion
Figure 1. A new approach for court exhibits relies not only on "points", but illustrates all three levels of ridge detail.

The first task is to acquire both the latent print and the inked print. When you acquire images either from a digital camera or from a film scanner, the images are not calibrated for one-to-one (1:1), life-size output. In contrast, if you are scanning images using a flatbed scanner, there is no depth of field involved and you specify both the size of the area to be scanned and the image resolution, so images acquired using this particular digital device are acquired as true, life-size images.

NOTE: It may be that you are scanning a photograph of an image that was enlarged, such as a latent fingerprint that was photographed and then printed as an 8 by 10 inch photograph. In this case, the image may or may not be enlarged by a specific scale, such as a 5 time enlargement, and so forth. Upon scanning this photograph, it would be necessary to rescale the image to print a life-size image. It is also necessary to have a scale (ruler) in an image to be calibrated.

In addition, please be advised that either when scanning an image or photographing an image with a digital camera, it is extremely important to acquire the image at the highest possible resolution in order to ensure the best possible image quality when printing the final court exhibit.
Noise and Inversion
Memory
Crime scene print for 5 seconds
Mask
Suspect print for 5 seconds
Same fingerprint or different?

Same  Different
Short Term Memory

- Experts (n=16) 67.88%
- Novices (n=42) 53.51%
Long Term Memory

- Experts (n=16): 54.19%
- Novices (n=40): 52.38%
Memory
Mask
2 or 60 seconds
Mask
Match?

Yes  No
Development of Expertise

Noise and Inversion

Memory

Deadline
mbthompson.com

Forensic Reasoning.com

Expertise and Evidence.com