

The style of a stranger: Identification expertise generalizes to coarser level categories

Rachel A. Searston¹ · Jason M. Tangen¹

Published online: 5 December 2016
© Psychonomic Society, Inc. 2016

Abstract Experience identifying visual objects and categories improves generalization within the same class (e.g., discriminating bird species improves transfer to new bird species), but does such perceptual expertise transfer to coarser category judgments? We tested whether fingerprint experts, who spend their days comparing pairs of prints and judging whether they were left by the same finger or two different fingers, can generalize their *finger* discrimination expertise to *people* more broadly. That is, can these experts identify prints from Jones's right thumb and prints from Jones's right index finger as instances of the same "Jones" category? Novices and experts were both sensitive to the *style* of a stranger's prints; despite lower levels of confidence, experts were significantly more sensitive to this style than novices. This expert advantage persisted even when we reduced the number of exemplars provided. Our results demonstrate that perceptual expertise can be flexible to upwards shifts in the level of specificity, suggesting a dynamic memory retrieval process.

Keywords Identification · Categorization · Family resemblance · Recognition · Perceptual expertise

Electronic supplementary material The online version of this article (doi:10.3758/s13423-016-1211-6) contains supplementary material, which is available to authorized users.

✉ Rachel A. Searston
rachel.searston@gmail.com

¹ School of Psychology, The University of Queensland, St Lucia, QLD, Australia 4072

Both common experience and laboratory research alike demonstrate that we (humans) are highly sensitive to the structure of visual categories. We see an unknown thing and draw on our previous experiences to decide what category it belongs to (Bruner, Goodnow, & Austin, 1956). As we begin to learn about the visual structure of a new category, each piece of information is equally informative, so its boundaries and "typical" instances are all unknown. Over time, we develop a tacit knowledge about how they tend to look and vary and rely on this knowledge to generalize new encounters (Brooks, 1978; Medin & Schaffer, 1978). As we accumulate more examples and verbal descriptions, the information that we once relied upon to distinguish between, for example, a sweet or dry wine, a songbird or wading bird, and a Monet and Picasso painting, is sharpened, which allows us to distinguish eventually between a German or Australian Riesling, house finches or purple finches, and a Blue or Rose period Picasso (Richler & Palmeri, 2014). This sharpening process is accompanied by an increase in the speed and accuracy of classification and better transfer to novel exemplars from the same category level (Tanaka, Curran, & Sheinberg, 2005; Tanaka & Taylor, 1991; Wong, Palmeri, & Gauthier, 2009).

By virtue of being an expert in a particular field, one can easily work within the constraints of that domain. For example, experience with identifying specific bird species transfers to the identification of novel bird species (Tanaka et al., 2005). Many experiments, however, have demonstrated the abrupt *limits* or inflexibility of expert performance when venturing too far afield: expertise with modern cars does not transfer to antique cars (Bukach, Phillips, & Gauthier, 2010), expertise with Tetris shapes does not transfer to non-Tetris shapes (Sims & Mayer, 2002), chess masters' memory for chess board configurations resembles a beginner's when the configurations are scrambled (Chase & Simon, 1973), and expert recognition is disrupted when images are presented in a novel orientation

(Diamond & Carey, 1986). Altering the appearance of the subject matter, even slightly from what an expert typically encounters, is enough to produce a significant drop in performance. However, there has been relatively little investigation of how experts generalize to coarser levels of specificity. Can the classification of Rieslings by region help people classify wine along broader dimensions, such as varietal, sweetness, or color? Does experience with specific bird species help people distinguish song birds from wading birds or birds from bats? Finally, does experience with discriminating Picasso paintings by period help people differentiate between different artists?

We turn to fingerprint identification as a testbed for measuring whether specific expertise can generalize to coarser level categories within a given domain. We used the domain of fingerprints, because it affords a rare sample of experts with extensive identification experience compared to a genuine novice control group. Fingerprint examiners spend their days visually comparing pairs of impressions side-by-side and judging whether they were left by the same or different fingers. These professionals also display hallmarks of genuine

perceptual expertise. They are impressively accurate compared with novices when discriminating prints by finger (Tangen, Thompson, & McCarthy, 2011; Ulery, Hicklin, Buscaglia, & Roberts, 2011; but see Ulery, Hicklin, Buscaglia, & Roberts, 2012, Dror & Rosenthal, 2008 and Dror & Cole, 2010, for issues of reliability and context effects in fingerprint examinations), and there is evidence to suggest that they rely on configural, holistic, or nonanalytic processes when matching prints (Busey & Vanderkolk, 2005; Thompson & Tangen, 2014).

In the following experiment, instead of judging whether two prints belong to the same or different *fingers*, our participants judge whether a series of five prints presented in a lineup belong to the same or different *person* (Fig. 1a). Can people identify prints from Jones’s right thumb, index, middle, ring, and little finger as instances of the same “Jones” category? Our expert participants have years of experience with matching prints from the same finger or different fingers but no experience with explicitly matching prints from the same person or different people. Novices, of course, have no

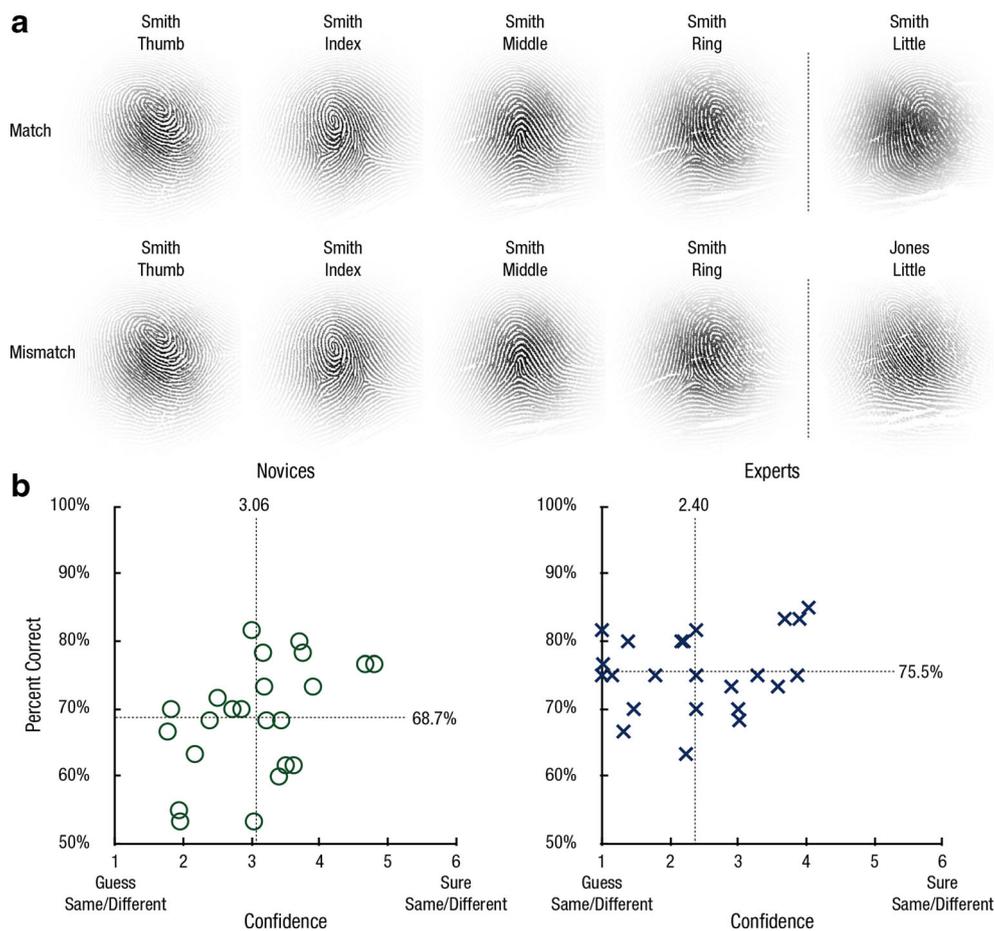


Fig. 1 Examples of matching and mismatching lineups (a). The results of the experiment (b) depicted as the mean percentage of correct responses for each individual along the y-axis and their mean confidence along the x-axis (novices are represented as green circles on

the left and experts as blue crosses on the right). The cross-hairs in each graph indicate the mean percentage of correct responses for each group (the horizontal line) relative to the mean confidence for the group (the vertical line)

experience with either task. Experts performing more accurately than novices would suggest a process akin to family resemblance categorization, where experts are drawing on their memory for how prints tend to vary, not just within and between fingers, but within and between people (Medin, Wattenmaker, & Hampson, 1987; Rosch & Mervis, 1975; Wittgenstein, 1953).

Method

Participants

Twenty-three qualified practicing fingerprint experts from four police agencies in Australia (The Australian Federal, New South Wales, Queensland, and Victoria Police) with an average of nine years of experience in matching fingerprints participated in the experiment, and 23 undergraduates from The University of Queensland—our novices—also participated for course credit. We recruited as many experts as possible, and an equal number of novices.

Stimuli

The stimuli were 10 fully rolled fingerprints collected from each of 60 individuals (600 fingerprints in total) and sourced from the Forensic Informatics Biometric Repository (Tangen & Thompson, n.d.). The prints were cropped to 600×600 pixels and we applied a Gaussian mask to each blurring the edges to isolate the structure of the prints. Sixty lineups (30 matching and 30 mismatching) were generated for each participant; each lineup always consisted of an impression from each of the five digits from the same hand type (Fig. 1a for an example of a matching and mismatching lineup). The lineups were sampled equally from the left and right hand and further partitioned equally as targets and distractors.

Specifically, for each participant, a random half of the identities in our image set were reserved for left-hand trials (i.e., the five prints for the left hand were used in the experiment and the right-hand prints were left out) and the other half of the identities were reserved for right-hand trials (i.e., the five prints from the right-hand were used and the left-hand prints were left out). Additionally, for each participant, a random half of the identities in each of the left- and right-hand trials were allocated for target lineups and the remaining half for distractor lineups. The digits in each lineup also were presented in random order on the screen, meaning that each digit type had a 1/5 chance of being a target on each of the match trials or a 1/5 chance of being replaced by a distractor on each of the mismatch trials. The distractor was always from the same hand and digit type but from another random individual. Targets and distractors were presented in a different random order for each participant.

Procedure

After reading an information sheet about the experiment and watching an instructional video, we presented 60 fingerprint lineups, 1 at a time. Participants were instructed to judge whether the fingerprint on the far right of each lineup (e.g., the little fingers in Fig. 1a) belonged to the same person or a different person from the first four. Participants were instructed about the nature of the lineups during the instructional video. That is, they were told that each lineup would consist of a thumb, index, middle, ring, and little fingerprint (in random order) and that the first four prints in each lineup were from the same person in each case.

Each fingerprint lineup remained on the screen until participants provided a response, and they indicated their judgments on a 12-point confidence rating scale ranging from 1 (sure different) to 12 (sure same); ratings of 1 through 6 indicated a “no match” decision and ratings 7 through 12 indicated a “match” decision. This forced-choice design provides measures of both discrimination ability and response bias in addition to the raw confidence scores (Vokey, Tangen, & Cole, 2009). Each participant judged 60 lineups: 30 “matching” and 30 “mismatching,” presented in random order.

Results

For each participant, we calculated the percentage of lineups responded to correctly over the 60 trials. We also calculated each participant’s absolute confidence scores over the 60 trials by converting each rating to a score out of 6 (e.g., ratings of 1 or 12 on the 12-point scale would each correspond to a confidence score of 6/6, and ratings of 6 or 7 would correspond to a confidence score of 1/6). See Fig. 1b, for the mean percentage of correct responses for each of the 23 novices and 23 experts relative to their mean confidence. There was no significant difference in the mean viewing time (i.e., time to respond) of novices (mean response time = 9.52 seconds) and experts (mean response time = 10 seconds), $t(44) = 0.53$, $p = 0.599$.

Accuracy

Both novices and experts performed quite well on this task. On average, novices correctly classified 68.7% of the lineups compared to experts who correctly classified 75.51% of the lineups. We computed the average discrimination ($A\neq$) and response bias ($B\equiv_D$) for novices and experts (see Vokey et al., 2009, for a similar analysis and discussion; see also Donaldson, 1992 for a discussion of $A\neq$ and $B\equiv_D$ as nonparametric measures of accuracy and response bias). Analyses of these measures allowed us to see whether the performance differences observed between novices and experts were due

to genuine differences in discrimination ability, or whether they were due to differing response thresholds (i.e., a tendency to say “match” or “no match” more often). A t test using novices’ and experts’ $A\prime$ scores revealed that experts ($Mean A\prime = .83$) were indeed significantly more accurate than novices ($Mean A\prime = .76$), $t(44) = 3.24$, $p = 0.002$, $d = 0.91$ (the same analysis using d_2 , a parametric measure of discrimination ability, revealed the same pattern of results). There was no significant difference in response bias between experts ($Mean B3_D = -0.02$) and novices ($Mean B3_D = -0.01$), $t(44) = 0.03$, $p = 0.764$, with neither group exhibiting a strong bias to overcall a particular outcome.

Confidence

Experts’ superior performance is particularly interesting in the context of their confidence ratings. Even though experts were significantly more accurate on this four-to-one matching task, they were also less confident in their judgments (mean confidence = 2.40/6) compared to novices (mean confidence = 3.06/6), $t(44) = 2.42$, $p = 0.020$, $d = 0.72$. Novices also displayed a significant (albeit weak) *positive* relationship between their confidence and percentage of correct responses, $r(21) = 0.47$, $p = 0.025$, but there was no significant relationship between confidence and accuracy for experts, $r(21) = 0.20$, $p = 0.351$.

Discussion

We have shown that people are sensitive to the *style* of a stranger. We asked participants to make a novel judgement about a set of fingerprints: given a lineup of four prints, did the fifth come from the same person or did it come from a different person? Half the participants—our novices—had no experience with fingerprints whatsoever. The other half—our experts—had several years of experience with fingerprints, but at a different, finer level of specificity. That is, experts compare fingerprints side-by-side and judge whether they were left by the same or different *fingers*; they have no explicit experience with matching *people*. Both groups generally performed the task well, but our experts were more accurate than our novices. Fingerprint experts were more sensitive to the style of a stranger than undergraduate novices, despite being less confident (see also the Supplemental Material available online for a replication and extension of this experiment where we reduced the number of impressions in each case from five to two and show that experts are more accurate than novices at distinguishing print pairs from the same person versus different people, even when pattern type cannot be relied on as a diagnostic cue). These data provide evidence that experts, with years of experience matching pairs of fingerprints, can

transfer this identification skill to categorizing prints from the same or different people, more broadly.

Anecdotally, when we asked experts about the basis for their decisions, some referred to vague similarities in the thickness of the ridges, or a similar ridge “flow” across the five prints, but the majority indicated they did not know for sure, expecting their performance to be quite poor. These results are consistent with models of automaticity that propose a shift from explicit, rule-based processing to more implicit, memory-retrieval processing with expertise (Logan, 1988). We revisited the confidence data from Tangen et al. (2011), where experts were much more confident in their decisions when matching prints from the same or different *fingers* (mean confidence = 5.09/6 compared to 2.40/6 in the current task). At this more familiar level of specificity, experts’ confidence was strongly and positively correlated with their average percentage of correct responses [$r(35) = 0.65$, $p < 0.001$], suggesting their metacognitive judgments aren’t well calibrated for identifying the limits and flexibility of their own expertise.

When matching *people* (versus *fingers*), it seems that participants are less aware of the dimensions that influence their accuracy and base their confidence on dimensions that have no bearing on their performance as a result. This explanation is consistent with recognition memory accounts of confidence-accuracy relations, which suggest that confidence ratings are made on the basis of different information to accuracy judgments (Busey, Tunnicliff, Loftus & Loftus, 2000). When identifying faces, for example, people can overestimate the impact of luminance on their accuracy—mistakenly believing that they are more accurate at identifying brighter faces when this is not the case (Busey et al., 2000).

Our primary interest, however, lies in the difference between experts and novices in how well they can distinguish between the same or different people. Experts were more accurate than novices, which suggests that subordinate level identification expertise can generalize to coarser level categorization judgments. Fingerprint examiners have no experience with explicitly classifying impressions of different fingers from the same person. However, it is likely that as they accumulate experience with generalizing from impression to impression, these experts develop a tacit sensitivity to the family-resemblances, covariant information, visual structure, or “style” among fingerprints, not just across instances of particular fingers, but across instances of people as well. From this perspective, the effects we have observed could be explained by experts accessing information that is distributed across their repository of prior instances. Indeed, previous work has demonstrated that fingerprint comparison judgments are influenced by similar past cases (Searston et al., 2015). Our results push this idea even further, illustrating flexibility in the way that

perceptual experts are able to retrieve and use their prior knowledge. This interpretation is consistent with some exemplar models of categorization, which assume that identification and categorization draw on the same underlying dimensions that are most optimal for performance of the task at hand (Nosofsky, 1986, 1987).

Perhaps our experts were simply more motivated than novices to perform well or entered the profession, because they have an “inherent visual ability” to match prints. Alternatively, prior work has demonstrated that experience with explicitly identifying and classifying visual categories typically results in improved performance with those specific categories (Tanaka et al., 2005; Wong et al., 2009), and our experiential account is consistent with this body of work. It also is unclear why experts would display reduced confidence in their judgments when discriminating people if the expertise effect we observed is purely a result of inherent ability. Similarly, novices’ rosier view of their own performance may have *increased* their motivation to perform relative to our experts. In academic writing, confidence in one’s ability is positively associated with measures of motivation as well as measures of performance (Pajares, 2003). From this view, if participants’ accuracy was influenced by their motivation to perform, it may have even dampened the expertise effect. Future studies that measure performance over time as learners gain experience in a particular domain will surely provide more insight into these issues of motivation and ability.

In this study, we set out to probe whether people who are already experts at making fine-grained visual discriminations, maintain an expert advantage when pushed outside of their usual level of specificity. Across two separate experiments, we have shown that they do. Our goal was not to make specific inferences about the particular information that fingerprint experts might rely on to produce the observed effects. Such information would certainly be useful in understanding the dimensions that are diagnostic of fingerprint matching, and others are making headway in this space (Busey & Parada, 2010; Busey, Yu, Wyatte, Vanderkolk, Parada & Akavipat, 2011). Our point in using “style” is to emphasize the information that remains latent in memory, so we have focussed less on fingerprint matching *per se*, and more on the flexibility of perceptual expertise: given extensive experience discriminating visual objects at a granular level, does this experience allow people to stretch across levels of specificity? In contrast to findings that expertise is static, inflexible, and highly task specific (see Lewandowsky & Thomas, 2009, for a review of some of these findings), our results provide an example of perceptual expertise that is more dynamic in nature, and flexible to upward shifts in the level of specificity.

Author Note This research was supported by grant No. LP120100063 from the Australian Research Council (ARC) to J. M. Tangen. The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorisation*. Hillsdale, NJ: Erlbaum.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.
- Bukach, C. M., Phillips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception & Psychophysics*, 72(7), 1865–1874. doi:10.3758/APP.72.7.1865
- Busey, T. A., & Parada, F. J. (2010). The nature of expertise in fingerprint examiners. *Psychonomic Bulletin & Review*, 17(2), 155–160. doi:10.3758/PBR.17.2.155
- Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence–accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, 7(1), 26–48.
- Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, 45, 431–448. doi:10.1016/j.visres.2004.08.021
- Busey, T., Yu, C., Wyatte, D., Vanderkolk, J. R., Parada, F. J., & Akavipat, R. (2011). Consistency and variability among latent print examiners as revealed by eye tracking methodologies. *Journal of Forensic Identification*, 61(1), 60–91. Available at: <https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=255267>
- Chase, W. G., & Simon, H. A. (1973). The mind’s eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic Press.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115(2), 107–117. doi:10.1037/0096-3445.115.2.107
- Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, 121(3), 275–277. doi:10.1037/0096-3445.121.3.275
- Dror, I. E., & Cole, S. A. (2010). The vision in “blind” justice: Expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, 17(2), 161–167. doi:10.3758/PBR.17.2.161
- Dror, I., & Rosenthal, R. (2008). Meta-analytically quantifying the reliability and biasability of forensic experts. *Journal of Forensic Sciences*, 53(4), 900–903. doi:10.1111/j.1556-4029.2008.00762.x
- Lewandowsky, S., & Thomas, J. L. (2009). Expertise: Acquisition, limitations, and control. *Reviews of Human Factors and Ergonomics*, 5(1), 140–165. doi:10.1518/155723409X448044
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4), 492. doi:10.1037/0033-295X.95.4.492
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207–238. doi:10.1037/0033-295X.85.3.207
- Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, 19(2), 242–279. doi:10.1016/0010-0285(87)90012-0
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115(1), 39–57. doi:10.1037/0096-3445.115.1.39
- Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition*, 13(1), 87–108. doi:10.1037/0278-7393.13.1.87
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*, 19(2), 139–158. doi:10.1080/10573560308222
- Richler, J. J., & Palmeri, T. J. (2014). Visual category learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1), 75–94. doi:10.1002/wcs.1268
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605. doi:10.1016/0010-0285(75)90024-9
- Searston, R. A., Tangen, J. M., & Eva, K. W. (2015). Putting bias into context: The role of familiarity in identification. *Law and Human Behavior*. Advance online publication. doi:10.1037/lhb0000154
- Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, 16(1), 97–115. doi:10.1002/acp.759
- Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world, perceptual expertise. *Psychological Science*, 16, 145–151. doi:10.1111/j.0956-7976.2005.00795.x
- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23, 457–482. doi:10.1016/0010-0285(91)90016-H
- Tangen, J. M., Thompson, M. B., & McCarthy, D. J. (2011). Identifying fingerprint expertise. *Psychological Science*, 22(8), 995–997. doi:10.1177/0956797611414729
- Thompson, M. B., & Tangen, J. M. (2014). The nature of expertise in fingerprint matching: Experts can do a lot with a little. *PLoS ONE*, 9(12), 1–23. doi:10.1371/journal.pone.0114759
- Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2011). Accuracy and reliability of forensic latent fingerprint decisions. *Proceedings of the National Academy of Sciences*, 108, 7733–7738. doi:10.1073/pnas.1018707108
- Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2012). Repeatability and reproducibility of decisions by latent fingerprint examiners. *PLoS ONE*, 7, e32800. doi:10.1371/journal.pone.0032800
- Vokey, J. R., Tangen, J. M., & Cole, S. A. (2009). On the preliminary psychophysics of fingerprint identification. *The Quarterly Journal of Experimental Psychology*, 62(5), 1023–1040. doi:10.1080/17470210802372987
- Wittgenstein, L. (2001 [1953]). *Philosophical investigations*. Blackwell Publishing. ISBN 0-631-23127-7.
- Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects. *Psychological Science*, 20(9), 1108–1117. doi:10.1111/j.1467-9280.2009.02430.x