Psychonomic Bulletin & Review
2010, 17 (2), 161-167
doi:10.3758/PBR.17.2.161

The vision in "blind" justice: Expert perception, judgment, and visual cognition in forensic pattern recognition

Itiel E. Dror
University College London, London, England
and Cognitive Consultants International Ltd., Southampton, England

and

Simon A. Cole
University of California, Irvine, California

Many forensic disciplines require experts to judge whether two complex patterns are sufficiently similar to conclude that both originate from the same source. Studies in this area have revealed that there are a number of factors that affect perception and judgment and that decisions are subjective and susceptible to extraneous influences (such as emotional context, expectation, and motivation). Some studies have shown that the same expert examiner, examining the same prints but within different contexts, may reach different and contradictory decisions. However, such effects are not always present; some examiners seem more susceptible to such influences than do others—especially when the pattern matching is "hard to call" and when the forensic experts are not aware that they are being observed in an experimental study. Studying forensic examiners can contribute to our understanding of expertise and decision making, as well as have implications for forensic science and other areas of expertise.

Many forensic identifications are based on matching a visual pattern left at a crime scene and one from a suspect. However, a large body of forensic trace evidence (such as shoe prints, firearms, tool marks, bloodstains, hair, finger and palm prints, bite marks, handwriting, ear prints, tire marks, etc.) lacks instrumental analysis. In these types of evidence, the "instruments" are, to a large extent, the human expert examiners themselves, who make judgments on the similarity of visual patterns (Cole, 2001; Haber & Haber, 2008; Mnookin, 2008). Such dependence on specialized human visual perception and judgment in expertise (Busey & Vanderkolk, 2005; Gauthier, Williams, Tarr, & Tanaka, 1998) is common in a wide range of domain experts, from radiologists to military fighter pilots (Berlin & Hendrix, 1998; Dror, Kosslyn, & Waag, 1993).

Fingerprint identification is among the most widely used forensic techniques. It is cognitively challenging because no two fingerprint impressions, even from the same source finger, are ever identical; along with intersource differences, there are also intrasource variations. Due to the elasticity of the skin, the pressure applied, the material on which the prints are left, the method of lifting the prints, and a variety of other factors, visual differences are always introduced, even in the best and most ideal cases. And in the real world of forensics, things are far from ideal; the marks left at a crime scene (the latent prints) usually are partial and distorted and include noise.

Thus, the role of expert fingerprint examiners is complex: They do not simply determine whether two images are "identical" but determine whether different images are sufficiently similar to conclude that they originate from the same source. This can be very challenging, because some intersource differences are extremely small, thus producing *look-alike* prints that are very similar but, nevertheless, originate from different sources (people). Performance levels are reduced, difficulty is increased, and potential problems of false positives arise as distractors become more similar to the target (Ashworth & Dror, 2000; Vokey, Tangen, & Cole, 2009). With the growing use of searchable databases, the potential for such error drastically increases, because it is more likely that one will find a look-alike nonmatch (Cole, 2005; Dror & Mnookin, 2010; Dror, Péron, Hind, & Charlton, 2005).

This is clearly an interesting area in visual cognition, combining issues in perceptual expertise, judgment, and decision making. It is particularly interesting because fingerprint experts have long been touted as infallible (Federal Bureau of Investigation, 1985). On the rare occasions in which errors are found, individual examiners are blamed, attributing the error to incompetence, negligence, or fraud, insisting that, in the hands of competent experts, errors are "virtually impossible" (Ashbaugh, 1994; Cole, 1998, 2005). Incompetence or deliberately made false expert identifications are of no interest to cognitive science.

I. E. Dror, i.dror@ucl.ac.uk

However, expert judgments that are sincere but nevertheless incorrect are a different story (Giannelli, 1997, 2006). Bias and other cognitive influences unconsciously affect hard-working, honest, and dedicated forensic experts, thus creeping in without the experts' awareness. This is a difficult and interesting problem, with generalizability across domains.

Errors committed by well-intentioned experts are more problematic and dangerous for at least three reasons. (1) Cognitive biases affect all examiners, not just "bad apples" (Thompson, 2008), and thus have a potential to impact forensic judgments across the board and to be relatively widespread. (2) Erroneous judgments of forensic experts are all the more persuasive in the legal context because the examiners believe them themselves. Even more than an honestly mistaken eyewitness, an honestly mistaken expert is the least culpable and thus, potentially, the most dangerous kind of witness that can testify in a legal proceeding. (3) Many individual examiners—and more worrisome, many forensic professional bodies (both in the U.S. and in Europe)—have been reluctant and resistant to acknowledge, accept, and take proper action to counter these biases. Risinger, Saks, Thompson, and Rosenthal (2002) contended that forensic science, as a practice, has historically dismissed cognitive bias by conceptualizing it as an ethical issue, to be overcome by moral discipline, rather than as an inherent cognitive issue, to be managed by bias-minimizing actions. This is especially of concern because evidentiary independence can be corrupted; Hasel and Kassin (2009) illustrated how forensic evidence and confession to crimes can be interlinked and interdependent.

A recent National Academy of Sciences report on forensic science (National Academy of Sciences, 2009) concluded that

> a body of research is required to establish the limits and measures of performance and to address the impact of sources of variability and potential bias. Such research is sorely needed, but it seems to be lacking in most of the forensic disciplines that rely on subjective assessments of matching characteristics. These disciplines need to develop rigorous protocols to guide these subjective interpretations and pursue equally rigorous research and evaluation programs. The development of such research programs can benefit significantly from other areas, notably from the large body of research on the evaluation of observer performance in diagnostic medicine and from the findings of cognitive psychology on the potential for bias and error in human observers. (p. 8)

In this article, we review the current knowledge of and research into forensic decision making and highlight the interesting cognitive issues to be researched, with an emphasis on latent fingerprint identification. Such research will have important implications not only for forensic disciplines, but also for other domains that rely on expert visual perception and interpretation. This area provides a great opportunity to research perception and cognitive processing of complex visual patterns, expertise, and decision making. It is hoped that research in this area will follow in the footsteps of the research on eyewitness testimony and face recognition, which has provided a better understanding of memory and facial information processing and has influenced and motivated reform in policing and the criminal justice system.

Criminal Case Studies

An analysis of publicly exposed cases of fingerprint misidentification showed that cognitive bias probably contributed to such errors (Cole, 2005). One reason for this conclusion was that more than half of the errors, rather than being exposed through reanalysis by a second expert, had been corroborated by additional analysts, sometimes as many as three. One possible explanation for this finding is that the "verifying" examiners might have been biased by the awareness of their colleague's conclusion. Even more suggestive, however, were four cases (18% of the total cases examined) in which even experts hired by the defense had corroborated what was later determined to be an erroneous identification. This suggests that cognitive bias was more powerful than the combination of the actual data and whatever motivation the experts might have had to interpret the evidence in a manner favorable to their own clients (Cole, 2005).

One such case was the Brandon Mayfield case, in which a Muslim attorney in Oregon was arrested and held for 2 weeks as a material witness in the 2004 Madrid terrorist bombing, on the basis of an erroneous match between his fingerprint and one found at the crime scene. This high-profile identification error undermined the claim that only "incompetent" practitioners make errors, because it involved the prestigious FBI latent fingerprint laboratory. The case is particularly suggestive because, not only did the initial FBI examiner make an erroneous attribution, but at least two additional FBI examiners assigned to check that examiner's work corroborated and verified the erroneous identification. And, after that, a highly regarded independent examiner appointed by the court to examine the evidence on behalf of Mayfield's defense also concluded that Mayfield was the source of the print. After Mayfield had spent 2 weeks in jail, the Spanish National Police matched the fingerprint to an Algerian national, Ouhnane Daoud. Subsequently, the FBI released Mayfield, apologized, and gave him $2,000,000 in a compensation settlement.

Of course, one possible reason that so many different experts reproduced the same error might be attributed to the stimuli themselves—for example, the prints were extremely similar. A report by the U.S. Justice Department's Office of the Inspector General (OIG) posits such an explanation, claiming that the similarity of Mayfield's and Daoud's friction ridge skin (the skin that produces finger, palm, and sole prints) was "an extremely unusual event" (Fine, 2006). However, the Madrid bomber print was searched against several large databases of prints, using automated fingerprint identification systems (AFISs). An AFIS produces a list of candidates that is then examined

by a human expert. But since the AFIS examines millions of prints, it is not a surprise that similar look-alike prints are found by the mere scope of the search (Cole, 2005; Dror & Mnookin, 2010; Dror et al., 2005). The similarity makes it easier for cognitive biases to take over and affect the identification decisions. As the bottom-up data become more challenging (e.g., because of coincidental similarity or other reasons, such as low quality and/or quantity of the data), top-down cognitive influences have more leeway to influence the identification process. In scenarios such as this, computer searches combined with cognitive bias may prove to be a pernicious mix.

The Mayfield case shows that second, third, and even defense experts may be influenced by the knowledge that another examiner has already concluded that two fingerprints come from the same source. Thus, subsequent examinations may be biased by the initial examination. An internal FBI report concluded that this was, indeed, the case and recommended that verifications be performed "blind" in "designated cases" (whatever those are; Stacey, 2004).

But as important, if not more important, is that the initial examination may also be biased. As the OIG report details, the FBI examiners reached their erroneous conclusions despite seemingly clear (with the benefit of hindsight bias, of course; Berlin, 2000; Harley, Carlsen, & Loftus, 2004) discrepancies between the crime scene print and Mayfield's. This happened because the expert examiners posited ad hoc explanations for the discrepancies (e.g., the discrepancies were "noise" because they derived from another fingerprint that was laid down previously on the same surface by another individual). In other words, they explained away and dismissed the discrepancies (which otherwise would not have permitted an identification). The reason they did not see the data for what they were was that they were already convinced that Mayfield was the source of the print, and this biased their perception and judgment of the actual data.

There are other interesting cases of erroneous judgment. Take, for example, the case of detective Shirley McKie from Scotland. In this case, a print left at a murder scene was identified as that of Shirley McKie, and after this identification was further verified by other experts, she was arrested. However, she contested the identification all along, and indeed many leading international fingerprint experts have categorically rejected that it is her print. Eventually, she received an apology for the erroneous identification and was awarded £750,000 in a compensation settlement (Association of Chief Police Officers in Scotland, 2000; Justice 1 Committee, Scottish Parliament, 2007; McKie & Russell, 2007). What is interesting about this case is that the print continues to be disputed. The issues of cognitive bias and psychological influences have been raised by both sides as a contributing factor to the errors of their opponents.

Research Studies

There has been just a handful of research studies into forensic decision making, most of which have been conducted only in the past few years (an exception is a study on microscopic hair comparison [Miller, 1987]). The few existing studies provide a mixed picture, with conflicting findings as to whether bias affects forensic experts or not (see, e.g., Hall & Player, 2008, who claimed that bias has no effect on forensic experts' decisions; for a rebuttal, see Dror, 2009; Saks, 2009). Latent print identification has been the main focus of forensic decision-making studies, because it has provided a stimulus-rich opportunity for research and because it is reputed to be among the most accurate and objective of the forensic disciplines.

If expert latent fingerprint examiners can be biased, then so, presumably, can firearm and toolmark examiners, shoe print examiners, bite mark examiners, document examiners, ear print examiners, and all the other forensic experts who heavily rely on human perception and judgment. In fact, one of the reasons researchers chose to investigate this forensic discipline was because of its long-standing reputation for accuracy and objectivity (Leo, 2001). If bias is present in this domain, we can quite confidently assume that it is present to the same degree, if not more so, in the other, less established domains. Even the robust and newer forensic domain of DNA profiling is susceptible to biased interpretation (Thompson, 2006).

Research studies into forensic decision making have varied along a number of dimensions. For example, some have defined error only as false alarms (incorrect identifications) and have not regarded misses (incorrect exclusions or insufficiency) as error, whereas other studies have regarded both as error. Some studies have been conducted by cognitive psychologists and academics (e.g., Vokey et al., 2009), whereas others have been conducted by the forensic examiners themselves (e.g., Hall & Player, 2008; see Risinger & Saks, 2003, for the dangers and pitfalls of conducting law-enforcement-sponsored research). In addition, practitioner-led research is potentially problematic in that there may be conflict of interest, pressures, and "role ambiguity," all deriving from the fact that the practitioners are both researchers and participants (Cole, 2008).

Another important difference in the research studies has been the use of different participant groups. Some studies have used general university students (mainly psychology), and some have specifically used trained forensic students. Other studies have used actual forensic experts; however, some studies used experts during covert data collection while the experts were doing routine casework and were not aware they were being studied, whereas some studies used experts who knew they were being tested.We review the literature by participant group because it captures two important elements in studying forensic expert decision making; the first is level of expertise (general students, forensic students, and experts), and the second is ecological validity (testing experts in experiments vs. covertly collecting data during real casework).

Studies in which students (general university students) were used. A recent study utilizing signal detection theory (SDT) reported three experiments whose aim was to examine the extent to which people can correctly

determine the source of fingerprints (Vokey et al., 2009). The results of psychophysics testing showed that participants (nonexpert, not even forensic students) could perform at levels characterized by Vokey et al. as "quite well"; however, performance varied as a function of similarity and the specific finger in question. The early call for the application of SDT to the study of forensics (Phillips, Saks, & Peterson, 2001) had never been systematically followed up until the Vokey et al. study.

To study whether emotional context can affect judgment on fingerprints, Dror et al. (2005) presented pairs of prints to psychology student participants, who were required to decide whether or not the two prints originated from the same source. The participants ($N = 27$) received 96 pairs of prints (48 of them were clearly from the same source or clearly not from the same source, and 48 were ambiguous). Before presenting the fingerprints, some participants were briefed about the background of the case. Half of the briefed participants were given a low-emotional context (i.e., the prints were lifted off stolen property), whereas the other briefed participants were given a high-emotional context (i.e., the prints were lifted off a weapon used in a brutal and unprovoked attack). Horrific photos from crime scenes were used to support the high-emotional-context experimental condition.

The results showed that when the stimuli were clear, the emotional context had no impact on the participants' deciding that the prints were from the same source (or not from the same source). However, when the prints were ambiguous, an identification was more likely to be made when they were presented in the high-emotional condition. Furthermore, for the ambiguous prints, in the control group that did not include any background information, 47% of the prints were judged to be from the same source. When the same prints were presented in the low-emotional context, 49% were judged to be from the same source, whereas in the high-emotional context, 58% were judged to be from the same source (for details, see Dror et al., 2005).

Studies in which forensic science students were used. In two experiments, Schiffer and Champod (2007) evaluated how forensic science students conduct pattern analysis. During the initial stage, the examiner is supposed to examine and note only important features and characteristics of the pattern of the print, prior to any comparison with another print pattern. This is commonly referred to as the *analysis* stage in what latent print examiners call the *ACE–V methodology* (A, analysis; C, comparison; E, evaluation; and V, verification).

In the first experiment, performance was examined before and after training in fingerprint identification (Schiffer & Champod, 2007). The finding showed that, posttraining, participants noticed more characteristics (minutiae) and there was higher consensus among the participants. In their second experiment, Schiffer and Champod investigated whether the amount of minutiae detected would be influenced when the comparison print was present and whether low- and high-emotional-context backgrounds (petty burglary vs. terrorism) would influence the analysis stage. They found that neither had an effect on the analysis of the prints.

Studies in which forensic experts who knew that they were being tested were used. An error rate of 1.041% was observed with experienced latent print examiners (Wertheim, Langenburg, & Moenssens, 2006). However, almost all the errors (1.007% of the 1.041% total error) were classified by the authors as *clerical* mistakes, and only 0.034% of errors were classified as and attributed to erroneous identification judgments. Cole (2006) questioned the classification of the clerical errors, and Vokey et al. (2009) characterized this study as "quite flawed," because "among other things, it lacked distractor test prints, so false positives could not be assessed" (p. 1024).

In another study, 43 expert examiners were given a set of six comparisons (Langenburg, Champod, & Wertheim, 2009). The participants were randomly assigned to one of three experimental conditions: a control group; a low-context group, which received, along with the prints, the conclusion reached by another examiner; and a high-context group, which received, along with the prints, the conclusion reached by an internationally respected and recognized expert. The study examined whether the contexts would affect the examiners' decisions. They also included a second control group of novices. The findings showed that the context did bias all the participants but that its effects were more noticeable for the novices than for the experts. Furthermore, the experts were more affected in the direction of inconclusive and exclusion, and less in the direction of identification.

A study of 70 fingerprint experts examined emotional effects (Hall & Player, 2008). Some participants were told that the fingerprints were related to forgery (the low-emotional-context condition), whereas other participants were told that they were related to murder (the high-emotional-context condition). The authors concluded that the emotional context did not have an effect on the final decisions reached by the examiners. However, many of the participants in the study did not even read the context scenario and, thus, could not have been affected by it (13 participants, 19% of the entire participant pool); nevertheless, the data from these participants were included in analyzing the effect of the scenario context (see Dror, 2009; Saks, 2009). Furthermore, although Hall and Player concluded that there was no effect of emotional context in their study, the experts taking part in the study reported that "the severity of the case had an effect on their analysis" (Hall & Player, 2008, p. 38). This contextual effect is further established and reflected by the fact that the magnitude of the effect was a function of the level of the emotional context: 52% of the examiners in the high-emotional-context condition said that they were affected by the context, but only 6% of the examiners in the low-emotional-context condition said that the context affected them.

Emotional and motivational factors have also been examined in a qualitative study (Charlton, Fraser-Mackenzie, & Dror, 2010). In this study, the emotions of the examiners

themselves and their motivations were explored. The findings showed general positive emotional effects associated with matching fingerprints and fear of making errors, and specific satisfaction related to catching criminals, which was most notable in solving high-profile, serious, or long-running cases.

Studies in which covert data collection of forensic experts working real cases was used. Collecting experimental data from experts is important in terms of ecological validity. However, even in field collection, data can be problematic, because the behavior of people is often affected under experimental conditions, or even when they are just observed. This is further and significantly exacerbated when the topic of study is expert performance. In such studies, experts are motivated to do well, and thus their attention and performance are not reflective of their routine performance during ordinary day-to-day work. Furthermore, the study of error is an especially problematic and sensitive issue to investigate, and any data collection must, therefore, not reveal the purpose of the study. Thus, for ecological validity, this line of experiments was done covertly, collecting data from fingerprint experts during their routine day-to-day work.

Only two empirical field studies of this sort have been conducted (Dror & Charlton, 2006; Dror, Charlton, & Péron, 2006; see Dror & Rosenthal, 2008, for their meta-analysis). The first study used the Mayfield case to provide a context that suggested that two similar prints were not from the same source. Five experienced expert latent print examiners, who at the time of the study had not seen the actual Mayfield prints, were used in the data collection. These experts were presented with a pair of prints that allegedly were those of Mayfield and the latent print of the Madrid bomber (thus giving them strong contextual cues that, although these prints looked alike, they were, in fact, from different sources). The fingerprint examiners were asked to analyze the prints and to focus on the actual prints while ignoring the context (that the prints were from the Mayfield case). However, rather than the Mayfield prints being presented, unbeknownst to the experts, they were actually presented with a pair of prints that they themselves had previously examined years ago in real criminal cases and had determined to be from the same source—that is, an identification. The materials were exactly the same as those they had analyzed in the past, except that the context was repackaged so as to suggest that the prints were *not* from the same source.

In this study, 4 of the 5 expert examiners contradicted their original conclusions. Three of them changed from *identification* to *exclusion*—consistent with the contextually biasing information. One of them changed from *identification* to *inconclusive*. Only 1 stuck to the original response (Dror & Charlton, 2006). Thus, most of the expert fingerprint examiners made decisions on the basis of the context, rather than on the basis of the actual information contained in the print. The experimental manipulation in this study enabled the researchers to scientifically assess *within*-expert performance. Thus, comparing experts with themselves under different conditions allows researchers to isolate variables, using the experts as their own control, and avoiding between-experts individual differences (Byrne & Eysenck, 1993).

A follow-up study (Dror & Charlton, 2006) with a different and larger sample replicated and expanded the previous study. The next study used 48 pairs of fingerprints and showed that expert fingerprint examiners could be biased in both directions (toward *individualization* as well as toward *exclusion*). Furthermore, this study established that bias can be induced by more ordinary contextual information (such as "the suspect confessed to the crime" or "the suspect has an alibi"); for discussions on the presumption of evidentiary independence and the contamination of one form of evidence by another, see Castelle and Loftus (2001), Hasel and Kassin (2009), and Loftus and Cole (2004). Finally, this study also demonstrated that although bias is more likely to occur with difficult and similar prints, it can also occur in less challenging cases. Dror and Rosenthal (2008) combined the data collected from Dror and Charlton (2006) and Dror et al. (2006) and subjected it to meta-analysis, so as to quantify the reliability and biasability of fingerprint experts' decision making.

Summary and Conclusions

Many disciplines in forensic science rely on examining partial and distorted trace evidence left at a crime scene. These visual patterns—be they tire marks, shoe prints, or bite marks (to name just a few)—are compared with *target* patterns from potential suspects. A number of studies have examined how well forensic experts from the established and most commonly used discipline of fingerprinting examine different visual patterns. Fingerprint identification relies on the ability of human examiners to make correct visual judgments. For the most part, matching fingerprints that are complete and of high clarity may well be a straightforward pattern-matching task. However, it can become quite complex and interesting when the comparison is made to a partial and distorted print (e.g., latent prints collected from a crime scene). In such cases, patterns do not match easily, and the expert examiner needs to determine whether they are *sufficiently similar* to conclude that both print patterns originated from the same source.

Research shows that various factors affect the perception and comparison of fingerprint patterns and that judgments and decisions are subjective and susceptible to influences. Across studies, it seems that extraneous information, such as emotional context, expectation, and motivation, affects decision making. In fact, studies have shown that the same expert examiners, examining the same prints but within different contexts, may reach different and contradictory decisions. However, some examiners seem more susceptible to such influences than do others. Furthermore, all examiners are influenced more by context and other extraneous information when the pattern matching is objectively "hard to call" and when they are not aware that they are taking part in an experimental study.

It is clear that experts in forensic identification have special abilities and cognitive processes specializing in pattern recognition but that these processes are also vul-

nerable to error (Busey & Dror, in press; Dror, in press). Forensic pattern recognition comprises issues in visual cognition, expertise, and decision making—par excellence, subject matters within the domain of cognitive science. Nevertheless, there has been only a handful of studies in this area, most of which have taken place only in the past few years. These studies have only begun to unravel the perceptual and cognitive processes involved in forensic pattern recognition. This area calls for additional research that will benefit the understanding of expert performance, as well as provide scientific insights into the real world of forensic identification.

AUTHOR NOTE

For more information related to this article, see www.cci-hq.com. This research was support by grants awarded to the first author by the National Institute of Justice, and interagency funding by the National Institute of Standards and Technology, Federal Bureau of Investigation, and Department of Defence (Contracts N41756-10-C-3307, 2009-DN-BX-K225, and 2009-DN-BX-K224), and a grant awarded to the second author by the National Science Foundation (Contract SES-0115305). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of any of the funding agencies. Correspondence concerning this article should be addressed to I. E. Dror, Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, England (e-mail: i.dror@ucl.ac.uk; www.cci-hq.com).

REFERENCES

Ashbaugh, D. R. (1994). The premise of friction ridge identification, clarity, and the identification process. *Journal of Forensic Identification*, **44**, 499-516.

Ashworth, A. R. S., & Dror, I. E. (2000). Object identification as a function of discriminability and learning presentations: The effect of stimulus similarity and canonical frame alignment on aircraft identification. *Journal of Experimental Psychology: Applied*, **6**, 148-157.

Association of Chief Police Officers in Scotland (2000). *Report of the scrutiny of the SCRO fingerprint bureau and structure of the Scottish fingerprint service.* Available at www.scottish.police.uk/main/campaigns/interim/interim.htm.

Berlin, L. (2000). Malpractice issues in radiology: Hindsight bias. *American Journal of Roentgenology*, **175**, 597-601.

Berlin, L., & Hendrix, R. W. (1998). Malpractice issues in radiology: Perceptual errors and negligence. *American Journal of Roentgenology*, **170**, 863-867.

Busey, T. A., & Dror, I. E. (in press). Special abilities and vulnerabilities in forensic expertise. In A. McRoberts (Ed.), *Fingerprint sourcebook*. Washington, DC: NIJ Press.

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, **45**, 431-448.

Byrne, A., & Eysenck, M. W. (1993). Individual differences in positive and negative interpretive biases. *Personality & Individual Differences*, **14**, 849-851.

Castelle, G., & Loftus, E. F. (2001). Misinformation and wrongful convictions. In S. D. Westervelt & J. A. Humphrey (Eds.), *Wrongly convicted: Perspectives on failed justice* (pp. 17-35). New Brunswick, NJ: Rutgers University Press.

Charlton, D., Fraser-Mackenzie, P., & Dror, I. E. (2010). Emotional experiences and motivating factors associated with fingerprint analysis. *Journal of Forensic Sciences*, **55**, 385-393.

Cole, S. A. (1998). Witnessing identification: Latent fingerprint evidence and expert knowledge. *Social Studies of Science*, **28**, 687-712.

Cole, S. A. (2001). *Suspect identities.* Cambridge, MA: Harvard University Press.

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

Cole, S. A. (2006). The prevalence and potential causes of wrongful conviction by fingerprint. *Golden Gate University Law Review*, **37**, 39-105.

Cole, S. A. (2008). Comment on "scientific validation of fingerprint evidence under Daubert." *Law, Probability & Risk*, **7**, 119-126.

Dror, I. E. (2009). On proper research and understanding of the interplay between bias and decision outcomes. *Forensic Science International*, **191**, e17-e18.

Dror, I. E. (in press). Paradoxical functional degradation in human expertise. In N. Kapur, A. Pascual-Leone, & V. S. Ramachandran (Eds.), *The paradoxical brain.* Cambridge: Cambridge University Press.

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, **56**, 600-616.

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to make erroneous identifications. *Forensic Science International*, **156**, 74-78.

Dror, I. E., Kosslyn, S. M., & Waag, W. (1993). Visual–spatial abilities of pilots. *Journal of Applied Psychology*, **78**, 763-773.

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 forensic science. *Law, Probability & Risk*, **9**, 47-67.

Dror, I. E., Péron, A. [E.], Hind, S., & Charlton, D. (2005). When emotions get the better of us: The effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology*, **19**, 799-809.

Dror, I. E., & Rosenthal, R. (2008). Meta-analytically quantifying the reliability and biasability of fingerprint experts' decision making. *Journal of Forensic Sciences*, **53**, 900-903.

Federal Bureau of Investigation (1985). *The science of fingerprints: Classification and uses*. Washington, DC: United States Government Printing Office.

Fine, G. A. (2006). A review of the FBI's handling of the Brandon Mayfield case. Washington, DC: U.S. Department of Justice Office of the Inspector General.

Gauthier, I., Williams, P., Tarr, M., & Tanaka, J. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, **38**, 2401-2428.

Giannelli, P. C. (1997). The abuse of scientific evidence in criminal cases: The need for independent crime laboratories. *Virginia Journal of Social Policy & the Law*, **4**, 439-478.

Giannelli, P. C. (2006). Forensic science. *Journal of Law, Medicine & Ethics*, **34**, 310-319.

Haber, L., & Haber, R. N. (2008). Scientific validation of fingerprint evidence under Daubert. *Law, Probability & Risk*, **7**, 87-109.

Hall, L. J., & Player, E. (2008). Will the introduction of an emotional context affect fingerprint analysis and decision-making? *Forensic Science International*, **181**, 36-39.

Harley, E. M., Carlsen, K. A., & Loftus, G. R. (2004). The "saw-it-all-along" effect: Demonstrations of visual hindsight bias. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 960-968.

Hasel, L. E., & Kassin, S. M. (2009). On the presumption of evidentiary independence: Can confessions corrupt eyewitness identifications? *Psychological Science*, **20**, 122-126.

Justice 1 Committee, Scottish Parliament (2007). *Inquiry into the Scottish criminal record office and Scottish fingerprint service.* Available at www.scottish.parliament.uk/business/committees/justice1/reports-07/j1r07-03-vol1-00.htm.

Langenburg, G., Champod, C., & Wertheim, P. (2009). Testing for potential contextual bias effects during the verification stage of the ACE-V methodology when conducting fingerprint comparisons. *Journal Forensic Sciences*, **54**, 571-582.

Leo, W. F. (2001). Fingerprint identification: Objective science or subjective opinion? *The Print*, **17**, 1-3.

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

McKie, I., & Russell, M. (2007). *Shirley McKie: The price of innocence*. Edinburgh: Birlinn.

Miller, L. S. (1987). Procedural bias in forensic science examinations of human hair. *Law & Human Behavior*, **11**, 157-163.

Mnookin, J. L. (2008). The validity of latent fingerprint identification:

Confessions of a fingerprinting moderate. *Law, Probability & Risk*, **7**, 127-141.

NATIONAL ACADEMY OF SCIENCES (2009). *Strengthening forensic science in the United States: A path forward*. Washington, DC: Author.

PHILLIPS, V. L., SAKS, M. J., & PETERSON, J. L. (2001). The application of signal detection theory to decision-making in forensic science. *Journal of Forensic Sciences*, **46**, 294-308.

RISINGER, D. M., & SAKS, M. J. (2003). Rationality, research and leviathan: Law enforcement-sponsored research and the criminal process. *Michigan State Law Review*, **4**, 1023-1050.

RISINGER, D. M., SAKS, M. J., THOMPSON, W. C., & ROSENTHAL, R. (2002). The Daubert/Kumho implications of observer effects in forensic science: Hidden problems of expectation and suggestion. *California Law Review*, **90**, 1-56.

SAKS, M. J. (2009). Concerning L. J. Hall and E. Player "Will the introduction of an emotional context affect fingerprint analysis and decision-making?" *Forensic Science International*, **191**, E-19.

SCHIFFER, B., & CHAMPOD, C. (2007). The potential (negative) influence of observational biases at the analysis stage of fingermark individualisation. *Forensic Science International*, **167**, 116-120.

STACEY, R. B. (2004). Report on the erroneous fingerprint individualization bombing case. *Journal of Forensic Identification*, **54**, 706-718.

THOMPSON, W. C. (2006). Tarnish on the "gold standard": Understanding recent problems in forensic DNA testing. *The Champion*, **30**, 10-16.

THOMPSON, W. C. (2008). Beyond bad apples: Analyzing the role of forensic science in wrongful convictions. *Southwestern Law Review*, **37**, 1027-1050.

VOKEY, J., TANGEN, J., & COLE, S. [A.] (2009). On the preliminary psychophysics of fingerprint identification. *Quarterly Journal of Experimental Psychology*, **62**, 1023-1040.

WERTHEIM, K., LANGENBURG, G., & MOENSSENS, A. (2006). A report of latent print examiner accuracy during comparison training exercises. *Journal of Forensic Identification*, **56**, 55-93.

(Manuscript received June 28, 2009;
revision accepted for publication October 16, 2009.)