Abstract

One study is presented which explores the biasing effects of irrelevant contextual information on a fingerprint matching task. Bias was introduced by providing the outcomes of a DNA test relating to each fictitious case under consideration. This was engineered to suggest either a match, no match, or an inconclusive outcome, and was thus either consistent, misleading or unbiased depending on the ground truth of each fingerprint pair. The results suggested that, when the difficulty of the fingerprint matching task was measurably increased, participants became more vulnerable to the biasing information. Under such conditions, when performance was good, misleading evidence lowered accuracy, and when performance was weaker, consistent evidence improved accuracy. As such, the results confirmed existing demonstrations of cognitive bias from contextual information in the fingerprint task. Moreover, by taking a process-based approach, it became possible to articulate the concerns, and the potential solutions, at each stage of the workflow. The results offer value for the forensic science community in extending the evidence-base regarding cognitive bias, and in articulating routes to improve the credibility of fingerprint decisions.

Keywords: fingerprint matching, cognitive bias, irrelevant contextual information

A biased opinion: Demonstration of cognitive bias

on a fingerprint matching task through knowledge of DNA test results.

Fingerprint analysis has been used for over a century as a means of comparing a suspect and a perpetrator within a police investigation. A ‘match’ or ‘identification’ decision means that the suspect may remain a ‘person of interest’ in a case. In contrast, a ‘no-match’ or ‘exclusion’ decision means that an innocent party may be cleared of charges or released from custody. Finally an ‘inconclusive’ decision means that the analyst does not have sufficient information to draw a definitive conclusion, and any charge or conviction would need to rest on alternative evidence. Fingerprint evidence has come to be relied upon in police investigations and in court proceedings, with some reports suggesting that it is infallible (FBI, 1985) and that errors are ‘virtually impossible’ (see Ashbaugh, 1994; Cole, 2005). Against this backdrop, concerns have recently been mounting regarding human error due to cognitive bias. Indeed, a growing body of research now suggests that these concerns are founded. Unfortunately, research to date has often been compromised by the use of small numbers of experts, or the use of university students who are naïve to the fingerprint task. The purpose of the present paper is to address this shortcoming by examining performance in a sample of students who have been trained in the task of fingerprint analysis according to expert commentaries. With such a population, a more powerful test of cognitive bias becomes possible, providing an opportunity to extend the evidence base on cognitive bias effects in the fingerprint matching task.

*The Value of Fingerprint Evidence*

The reliance on fingerprints as a means of identification rests on two principles: persistence and uniqueness (Pankanti, Prabhakar, & Jain, 2002). The principle of persistence assumes that an individual’s basic fingerprint pattern is determined whilst in-utero, and (barring extreme measures) will not change over their lifetime. Alongside this, the principle of uniqueness assumes that nature never repeats itself and that no two individuals share the same fingerprints. Indeed, whilst individuals may have similar prints, even identical twins do not have identical prints. These two principles mean that fingerprint evidence has been viewed as a robust means of identification, and courts of law around the world have deemed fingerprint evidence admissible.

In recent years, however, concerns have begun to emerge regarding the reliability of the conclusions that are drawn. Once such case involved the investigation into the Madrid Bomber who detonated 10 bombs on trains within Madrid in 2004. The latent fingerprint found on the bag of detonators was examined, and the automated fingerprint identification system (AFIS) produced a list of possible matches. Brandon Mayfield, a US attorney, was on this list, and was held in custody for two weeks on the basis of a ‘match’ decision between his print and the latent print. Three FBI analysts confirmed the initial match decision. The error in this case was that Mayfield’s fingerprint, whilst similar, was not a match to the latent print. In fact, Mayfield had never been to Spain and did not even own a passport. Nevertheless, the knowledge of Mayfield’s Muslim beliefs, and his involvement in defending a known terrorist, may have affected the initial decision, and the lack of a blind verification procedure may have contributed to the confirmation of the initial error. Spanish police later arrested Ouhnane Daoud whose fingerprint was a better match to the latent print, and the FBI were forced to admit their mistake (see OIG report into the FBI Mayfield Case, OIG, 2006). Cases such as that of the Madrid Bomber have revealed that fingerprint analysis can be fallible. As a result, the National Academic of Sciences (NAS, 2009), and more recently, the UK Forensic Science Regulator (2015), have highlighted the danger of cognitive bias across the forensic sciences, and both have called for a substantial research effort to define the conditions under which this bias may emerge, and solutions to mitigate the resultant risks.

Cognitive bias arises when performance on a task is affected by factors unconnected to the task itself. It is demonstrated when someone is presented with contextual information that affects or alters the perception or decisions that they may reach (see Dror, in press, for a useful discussion of seven types of bias). For example, people perceive there to be a greater similarity between the facial composite of a perpetrator and a suspect if they are led to believe that the suspect is guilty (Charman, Gregory & Carlucci, 2009). Similarly, people are more likely to judge a suspect fingerprint to be a ‘match’ to the latent print from a crime scene if the suspect fits the stereotypic demographic for the particular crime (Smalarz, Madon, Yang, Guyll & Buck, 2016). These effects are assumed to arise because our preconceptions, expectations, or beliefs of guilt can affect the perception of, or the weighting assigned to, the evidence itself (Tversky & Kahneman, 1974; Kassin, Dror & Kukucka, 2013) providing a ‘fast and frugal’ heuristic or rule of thumb to guide decision-making when time or resources are limited (Gigerenzer, 2004). Saks, Risinger, Rosenthal & Thompson (2003) suggested that this creates an ‘investigative echo chamber’ in which consistent facts or conclusions seem to reverberate and grow stronger than they really are. The result can be a ‘bias snowball effect’ (Dror, 2012) in which expectations may incorrectly guide interpretations.

In the context of fingerprint analysis, cognitive bias may be demonstrated when the outcome of the matching task is affected by details that may be relevant to the case but are irrelevant to the fingerprints themselves (see Stammers & Bunn, 2015). These may include the knowledge held by others, details of the case, crime-scene images, or the results of other forensic tests. Dror, Kassin, and Kukucka (2013) note the particular vulnerability of fingerprint analysts to cognitive bias given that the process is relatively ill-defined, and the matching task itself can be very difficult. Indeed, the available evidence suggests that cognitive bias may be demonstrated at every stage of the fingerprint task.

*Demonstration of Cognitive Bias at Each Stage of Fingerprint Analysis*

*Analysis:* Fingerprint analysis proceeds through a process of *Analysis, Comparison, Evaluation* and then *Verification* (the ACE-V method). The first *Analysis* stage involves a decision as to whether the latent fingerprint is of sufficient quality for a comparison to be undertaken. If suitable, the fingerprint characteristics, or minutiae, are then identified and logged. Whilst some authors have shown this stage to be relatively robust and free from bias (Schiffer & Champod, 2007), others have demonstrated that decisions can be influenced by extraneous details. Specifically, Fraser-Mackenzie, Dror, and Wertheim (2013) have shown that the decision regarding print suitability, as provided by 24 experts, was affected both by the presence of a comparison print, and by knowledge of the decisions of other analysts. Similarly, Earwaker, Morgan, Harris & Hall (2015) showed that the decision regarding print suitability was affected by the severity of the case. In their study, print quality was first assessed independently by three fingerprint experts, and this provided a benchmark against which to compare the decisions of the participants in the study. The participant then judged the print suitability of a set of 20 prints presented in the context of a serious crime (murder) and a less serious volume crime (theft from a car). The findings suggested that the participants judged more ‘poor quality’ prints to be suitable for analysis when the case was serious than when less serious. Similarly, they inappropriately discarded more ‘prints of sufficient quality’ when the case was less serious. Both findings indicated a change to the decision regarding print quality to the very same prints as a result of the crime context they were perceived in.

Following the determination of suitability, Dror, Champod, Langenburg, Charlton, Hunt, and Rosenthal (2011) showed bias in the detection and recording of minutiae. In particular, they noted that analysis of the latent print was affected by the presence of a comparison print such that fewer minutiae were recorded when the latent was analysed in the presence of a comparison print than when analysed solo. Both the judgement of print suitability, and the identification of minutiae, should be tasks that depend solely on the latent print itself. In this regard, the fact that a performance is affected at all by a comparison print, the crime type, or the views of others, may be taken as a demonstration of cognitive bias.

*Comparison and Evaluation:* The literature has also reported cognitive bias at the comparison and evaluation stages of the fingerprint matching task. For example, Dror, Péron, Hind, and Charlton (2005) tested 27 student participants across 96 trials and demonstrated clear cognitive bias when irrelevant emotional contextual details were introduced. Dror et al. (2005) achieved this by manipulating the amount of emotional information that participants received about a case. In one quarter of the trials, they received no information at all and these trials represented the control condition. In a further quarter of trials, participants were presented with fingerprints in the context of a low-emotion crime such as a burglary, along with photographs of the items stolen. In another quarter of trials, participants were presented with fingerprints in the context of high-emotion crime such as a personal attack, with graphic photographs showing the injuries sustained by the victim. In the final quarter of trials, participants were presented with fingerprints in the context of the high-emotion crime, with case details and photographs accompanied by subliminal messages which took the form of the words ‘guilty’ and ‘same’ presented alongside the fingerprints but below the level of conscious awareness . The results suggested that the rate of ‘match’ decisions increased as the emotional information became more powerful. This effect only emerged if the fingerprint pair was difficult to judge as participants were more likely to be swayed by accompanying information when the decision was not clear-cut. Under such circumstances, the students tested showed a clear effect of cognitive bias from the accompanying contextual details.

A similar result was demonstrated by Searston, Tangen and Eva (2015, Experiment 1) who asked student participants to determine whether two fingerprints matched or not. Prints were presented in the context of fictitious cases that followed the manipulation of Dror et al. (2005) in terms of severity. Within this study, however, a distinction was drawn between the capacity to discriminate between matching and non-matching pairs of prints, and the response criterion to say ‘match’ or ‘no match’. The results interestingly revealed no difference in discrimination as a result of crime severity. However, there was a significant difference in the participants’ response criterion in that they were swayed to say ‘match’ more often in the more severe case. This finding provides valuable insight into the potential mechanisms underlying a shift in performance as case severity varies.

A similar demonstration of cognitive bias at the comparison stage has also been revealed in fingerprint experts rather than students. Dror and Charlton (2006) tested the performance of six experts on eight fingerprint comparisons. Four of these represented a test of consistency of decision making over time and thus did not involve bias. However, the remaining four comparisons involved the presentation of additional contextual information designed to encourage either a ‘match’ decision (the suspect has confessed) or a ‘no-match’ decision (the suspect has an alibi). Importantly, in each case, the contextual information ran counter to the ground truth of the trial as judged by the expert previously. Thus, the experiment was engineered to see whether the experts would change their minds. In this regard, the results showed that three of the six experts gave decisions which yielded to the direction of the contextual information. This suggested that the presence of contextual information can affect decision-making even in experts.

In contrast to these results, Hall and Player (2008) tested a large population of 70 experts but found no clear evidence of cognitive bias. Their manipulation relied on the presentation of a fingerprint pair in the context of either a low-emotional crime (such as a forgery) or a high-emotional crime (such as a murder). The results suggested that whilst the experts felt that their decision-making had been affected by the context, their actual performance was consistent across the two conditions suggesting no bias.

These mixed results may in part be attributable to the very different numbers of experts tested by Dror and colleagues (Dror et al., 2005; Dror & Charlton, 2006), Searston et al. (2015), and by Hall and Player (2008), and thus the different levels of statistical power represented in each study. However, the results of Hall and Player (2008) have also been questioned in terms of whether the contextual manipulation was noticed by all participants (Dror, 2009), and in terms of whether the participants could have talked to each other to reveal the manipulation and thus weaken its effect (Saks, 2009). Additionally, results have tended to be judged as more reliable when they came from experts who were unaware of the test (through insertion of trials into their normal caseload) rather than from experts who may have responded defensively or cautiously within a conscious test situation (Dror, 2009; Saks, 2009). Nevertheless, most demonstrations of cognitive bias due to contextual information have come from the same lab, and have used naïve students or small numbers of experts. In this respect, further demonstration of these bias effects could only strengthen the evidence-base.

*Verification:* Cognitive bias through the presentation of contextual information has also been shown when considering the final *verification* stage of fingerprint analysis. During this stage, a second analyst is asked to verify the decision of a primary decision-maker. Thus, this stage acts as a quality control mechanism. However, three studies exist which suggest that the second analyst can be affected by knowledge of the primary decision. For instance, Langenburg (2009) tested 6 experts in a complex design which involved initial decisions and then verifications to 30 trials. Importantly, the verifications were conducted by analysts who knew the original decision and knew the original analyst. The results suggested that all initial false positive errors (n = 9 in total) were detected at the verification stage. In contrast, however, all false negative errors (n = 6 in total) were repeated suggesting that the second analyst could be influenced by the original decisions. The small number of experts, however, meant that the results were interpreted with caution.

A far larger analysis was reported by Langenburg, Champod, and Wertheim (2009). They tested 43 experts and 86 novices, each of whom was asked to verify 6 fingerprint decisions (3 matches and 3 non-matches). One third of the participants in each group performed their task ‘blind’ to the original decisions, and these participants thus represented the unbiased (control) group. A further third of the participants were provided with the original decisions which were attributed to ‘trained fingerprint examiners’. The final third of the participants were provided with the original decisions which were attributed to ‘internationally recognized experts’. The results suggested that the novices were substantially affected by knowledge of the original decisions, as may be expected. Worryingly, however, the expert participants were also affected by knowledge of the original decisions such that there was a substantially higher rate of ‘inconclusive’ responses from the experts when the original decision had been similarly cautious. This suggested that performance on the verification task can be affected by factors beyond the task itself, again illustrating a cognitive bias.

Finally, Dror Charlton, and Péron (2006) revealed a similar influence of cognitive bias at the verification stage. They asked 5 experts to re-examine a single pair of prints which had been examined previously within their caseload. All experts had judged this pair to be a ‘match’, however, it was presented to them again as if it related to the Madrid bomber case. This case is well-known to have involved a wrongful match; hence the contextual information in this study was designed to sway the experts into thinking that the (matching) fingerprint pair was in fact not a match. Worryingly, the results indicated that 4/5 of the experts changed their decision in the face of the contextual information. Of these, three experts judged the pair to be ‘inconclusive’ and one now judged the pair as ‘not a match’. Only one expert gave the same, correct, match decision as before. This study was notable through the testing of experts who undertook the task without awareness within their normal caseload. Under such conditions, knowledge of the original decision again appeared to have a powerful biasing effect on performance at this verification stage.

*The Present Study*

The present study further examines the evidence for cognitive bias in the fingerprint matching task. Value is added here through the use of a large population of participants who, whilst not expert, have been trained through the use of expert commentaries and are thus not naïve to the task. Each participant completed a total of 72 trials, providing statistical power to the current test. Specifically, the strong number of participants and trials enabled an assessment of performance that was less susceptible to any inadvertent testing or response errors, and was thus more representative of participant performance. Moreover, it enabled the use of more powerful statistical analyses to investigate the predictions at hand. As such, this study sits alongside those studies using smaller numbers of experts who completed relatively few trials (such as Dror & Charlton’s (2006) study of six experts across 8 trials, or Dror et al.’s (2006) study of 5 experts on one critical trial) and seeks to replicate their findings.

The present study took care to examine performance under cognitively easy and cognitively challenging conditions given previous demonstrations of bias only when the task was difficult. Here, however, task demands were manipulated not through changing the fingerprints themselves (where the change in difficulty is hard to quantify) but through changing the time available to inspect them (where the change can be defined with certainty).

Finally, value is offered in the present study through examining cognitive bias through the manipulation of a powerful yet plausible detail – the outcome of a DNA test related to the case. This manipulation builds on a prior study in the field of forensic anthropology which demonstrated that DNA results provided a powerful contextual bias when determining the sex, ancestry and age at death from skeletal remains (Nakhaeizadeh, Dror & Morgan, 2014). However, the effect of contextual information from DNA results has not yet been investigated in the context of a fingerprint matching task. Its use here is preferred over previous manipulations for three principal reasons. First, because DNA evidence has been viewed as the gold standard of forensic science, it may be judged as particularly believable or trustworthy by participants. Second, the evaluation of DNA evidence is less likely to be affected by the personality or emotionality of the participants in the fingerprint task compared to an emotional manipulation (Dror et al., 2005, Earwaker et al., 2015; Hall & Player, 2008; Searston et al. 2015). Third, DNA evidence may be judged as more reliable compared to suspect-related details such as a confession/alibi (Dror & Charlton, 2006) as suspect statements may not always be true.

With these factors in mind, participants completed a fingerprint matching task to ‘matching’ and ‘non-matching’ pairs of prints under conditions in which the DNA results either encouraged a ‘match’ decision, a ‘no-match’ decision, or provided no bias. On the basis of previous results, it was predicted that performance would be swayed in the direction of the contextual information, especially when task demands were increased through time pressure. More specifically, knowledge of a DNA match was predicted to improve accuracy on ‘matching’ trials, and decrease accuracy on ‘non-matching’ trials, whilst knowledge of a DNA no-match was predicted to have the reverse effects, relative to the unbiased condition.

Method

*Design*

Participants took part in a fingerprint matching task. A 2 x 2 x 3 mixed design was adopted in which time pressure (high, low) was varied between-participants. In addition, both trial type (‘matching’, ‘non-matching’) and contextual information (DNA match, DNA unclear, DNA does not match) were varied within-participants. Accuracy of fingerprint decision was recorded (compared to ground truth), and represented the dependent variable.

*Participants*

A total of 48 participants (40 females, 8 males) took part on a volunteer basis or in return for course credit. Ages ranged between 18 and 59 years (*M* = 20.56 years, *SD* = 5.31) and all participants confirmed normal, or corrected-to-normal, vision. In addition, all participants declared themselves as novices in the task of fingerprint analysis, and had no more than a rudimentary awareness of the basic fingerprint patterns at the start of the experiment.

Participants were randomly assigned to one of two groups which differed only in the time permitted for each fingerprint matching trial. Those in the low time pressure group had an unlimited amount of time to study the fingerprints. In contrast, those in the high time pressure group had only 2 seconds[[1]](#footnote-1) to study the fingerprints before they were removed from view, this being half of the viewing time used by participants in a previous study (Stevenage & Pitfield, 2016). A similar exposure duration of just 2 seconds was used by Thompson and Tangen as a way of manipulating task demands in experts and novices (Thompson & Tangen, 2014, Experiment 4).

*Materials*

Fingerprint stimuli were drawn from the FVC2004 DB1 Database (http://bias.csr.unibo.it/fvc2004/databases.asp), which consisted of 8 images of a single fingerprint from 100 individuals. These were collected under supervised laboratory conditions, across three separate occasions separated by at least two weeks. Fingerprints were recorded using a V300 Crossmatch sensor yielding images of 640 x 480 pixels with a resolution of 500 dpi. The sensor was not systematically cleaned between participants with the result that images were not necessarily pristine. Images within this database were discarded if blank or substantially smudged. However, fingerprint image quality was purposely varied through instruction to place the finger at slightly different angles and to apply more or less pressure. In this way, a distinction could be drawn between low quality prints, and high quality prints, which was useful for the current purposes.

*Selection of stimuli:* The 100 individuals were previously classified into fingerprint pattern types yielding three dominant groups corresponding to plain whorls (n = 27), radial loops (n = 31), and ulnar loops (n = 35). A total of 12 individuals were selected from each of the three pattern types (yielding 36 in total) and these represented the target prints. Each target print was depicted by two images representing a good quality image (simulating the ‘suspect print’ obtained from the custody suite or the 10-card), and a relatively poor quality or partial print (simulating the ‘latent print’ obtained from the crime scene). These judgements were made by the experimenters on the basis of the clarity of the ridge detail, and the completeness of the print itself (see Figure 1 for example stimuli).

(Please insert Figure 1 about here)

In addition, 6 individuals were selected from each of the 3 pattern types (yielding 18 in total), and these represented the foil prints for use on ‘non-matching’ trials.

Finally, 12 individuals were selected spanning the 3 pattern types and these represented practice prints. Four of these were used to construct ‘matching’ trials, and thus two images of each print were obtained as above. The remaining 8 stimuli represented the 4 targets and the 4 foils in 4 ‘non-matching’ trials, with care taken to match the fingerprint pattern across target and foil.

*Construction of fingerprint trials:* The fingerprint matching task consisted of a practice phase of 8 trials (4 ‘matching’, 4 ‘non-matching’) and a main phase of 72 trials (36 ‘matching’, 36 ‘non-matching’). In constructing the main phase of the task, each of the 36 target prints was presented twice, once in a ‘matching’ trial, and once in a ‘non-matching’ trial. For ‘matching’ trials, care was taken to ensure that identical images were never presented. Instead, the good quality target print (the suspect print) was paired with the corresponding poor quality target print (the latent print). This selection of stimuli ensured that (i) the task was not too trivial, (ii) the task more closely approximated that in the real world, and (iii) the use of simple image matching strategies on the part of the participant was minimised. For ‘non-matching’ trials, care was taken to ensure that the foil print was always of the same pattern type (whorl, radial loop, ulnar loop) as the target print, again ensuring that the task was not too trivial.

*Introduction of contextual information:* In order to present the contextual information, all main trials were accompanied by a statement summarising the results of a DNA test pertaining to the case for which each pair of fingerprints was being examined. One third of trials (half ‘matching’, half ‘non-matching’) were accompanied by contextual information indicating a DNA match between perpetrator and suspect. A further third of trials were accompanied by contextual information indicating no DNA match between perpetrator and suspect, and the final third of trials were accompanied by unbiased contextual information indicating insufficient evidence either way. As such, the accuracy of the contextual information (consistent, neutral, misleading) depended on the nature of the trial (‘matching’, ‘non-matching’) such that information indicating a ‘DNA match’ would be consistent in a ‘matching’ trial but would be misleading in a ‘non-matching’ trial, and so on. Care was taken to counterbalance the allocation of individual fingerprint pairs to each of these three contextual conditions across participants.

Trials were presented, and data were recorded, via Superlab 4.5 running on an HP Pavilion G Series laptop running Windows 10, with a 15” colour monitor and a screen resolution of 1366 x 768 pixels. Within this environment, each print measured approximately 8cm high x 5.3cm wide.

*Procedure*

Participants were tested individually within a quiet environment. Following the provision of informed consent, all participants engaged in a two-part procedure. In the first part, they all received training on the process used to determine whether two fingerprints represented a match (Data in Brief, submitted). This training was developed following interviews with 12 experienced analysts at a nearby UK Fingerprint Bureau (Stevenage & Pitfield, 2016), and it described both the information that analysts looked for in a fingerprint, and the process used when making a comparison. As such, participants were taught about the three basic fingerprint patterns, and about the appearance and value of deltas, bifurcations, ridge endings, pores, scars and creases with appropriate reference to these details within a fingerprint image. Equally, they were taught about the ACE-V method of analysis, comparison, evaluation and verification used by fingerprint experts when comparing one print with another. This stressed the primary need for the two prints to have the same overall fingerprint pattern if analysis was to proceed. It also stressed the need to assess the latent print, and then look for corresponding features in the suspect print. Given that the experts had not articulated a quantitative standard when determining a ‘match’ decision, no mention of any such standard was made during training. Instead, the training described the match decision as a judgement made by the analyst when weighing the evidence regarding points of correspondence during the evaluation phase.

The training tool took the form of a set of slides, annotated with quotes provided by the experts themselves so as to avoid inadvertent misinterpretation of their comments. The experts verified the accuracy and usefulness of the training tool prior to use. One of the experimenters talked through the slides with each participant, providing illustrations, clarifications and the opportunity for questions. There were no training trials at this stage. This training phase lasted no more than 20 minutes, and accuracy levels of 85.2% when given unlimited time and no contextual information suggested that this training was effective[[2]](#footnote-2).

In the second part of the procedure, all participants undertook a computer-based fingerprint matching task. This began with 8 practice trials, which enabled participants to orient to the response keys. Each practice trial was initiated by the presentation of two fingerprint images, displayed side by side on the computer screen. In addition, a prompt question “’Match’ or ‘Not a Match’?” was displayed below the images. Participants were asked to press ‘M’ if they thought the prints came from the same person and thus were ‘matching’, and they were asked to press ‘N’ if they thought the prints can from different people and thus were ‘non-matching’. On these practice trials, participants were allowed unlimited time to inspect the fingerprint images, however, they were encouraged to respond as quickly but as accurately as possible given the real-world consequences of a wrong decision. Feedback was provided on their answers whether correct or incorrect via an onscreen message accompanied by an audible beep in the case of an error. Poor performance at this stage resulted in the experimenter offering further assistance, with reference to the fingerprint images used in the training tool, and this reminded the participants of the points of correspondence to look for, and of the implications of an incorrect response in the real world.

Following the practice trials, and the opportunity to ask any clarifying questions, participants completed a set of 72 experimental trials, consisting of 36 ‘matching’ trials and 36 ‘non-matching’ trials presented in a random order. These differed from the practice trials in three important respects. First, for all participants, feedback was no longer provided. Second, the time allowed for fingerprint inspection differed across participants. For those in the ‘low time pressure’ condition, the fingerprints and prompt question remained in view until participant response. In this respect, experimental trials mirrored the practice trials. However, for those in the ‘high time pressure’ condition, the fingerprints were shown for only 2 seconds. After this period, a brief beep alerted the participants to the fact that their exposure time was over. The fingerprints were removed from view, and only the prompt question remained on screen until participant response. For all participants, however, the instruction remained in place to respond as quickly but as accurately as possible.

The third and most important difference between practice and experimental trials was that fictitious information was provided for all participants about the case to which each pair of fingerprints pertained. This served as a means to test for cognitive bias as a result of the delivery of contextual information. In this regard, participants were informed that the fingerprints were being considered within a murder enquiry for which DNA had been recovered from the scene, and a DNA test had already been conducted. For one third of the trials (12 ‘matching’, 12 ‘non-matching’), participants were informed that the DNA test had revealed a MATCH between the suspect and the perpetrator, thus providing a suggestion that the fingerprints may also be considered to represent a match. For a further third of trials (12 ‘matching’, 12 ‘non-matching’), participants were informed that the DNA test had revealed NO MATCH between the suspect and the perpetrator, thus providing a suggestion that the fingerprints may also not represent a match. For a final third of trials (12 ‘matching’, 12 ‘non-matching’), participants were informed that the DNA result had been UNCLEAR and that there was insufficient evidence to determine whether there was a match between suspect and perpetrator. This represented an unbiased or control condition in as much as there was no directional expectation implied.

In order to mitigate against concerns that participants may become ‘blind’ to the contextual information as the study wore on, the DNA MATCH message was shown in green, the DNA NO MATCH message was shown in red, and the DNA UNCLEAR message was shown in black. Thus, the direction of the suggestion was clear even if the words were no longer read. Additionally, a tagline summarising the contextual information was provided above the fingerprints themselves, stating ‘DNA is a match’, ‘DNA is not a match’, or ‘DNA is inconclusive’, in green, red, or black typeface respectively. Whilst the introduction of colour may have produced an emotional response in the participants, we selected the colours given their potential implicit association with actions (Green for ‘go’ or ‘yes’; Red for ‘stop’ or ‘no’). As such, it was hoped that colour would reinforce the directionality of the contextual information in each condition.

With these elements in place, participants completed the sequence of 72 trials in a randomised sequence, with a break at the half way point to minimise fatigue. The entire study lasted 35-40 minutes, after which participants were thanked and debriefed.

*Results*

In order to ensure that participants had attended to the training, performance was recorded during the practice trials. Examination of performance on these trials resulted in no participant data being dropped on the basis of poor performance. Consequently, the data may be taken to reflect performance of a trained though non-expert group.

The main task was a simultaneous matching task to pairs of fingerprints. Half the trials represented matching pairs and the other half represented non-matching pairs. Given the use of a 2 alternate, forced-choice task, it was possible to compute measures of sensitivity of discrimination (d’) and bias (C) from the accuracy scores according to a signal detection framework (Green & Swets, 1966). This has the benefit of balancing the number of correct ‘match’ decisions to matching trials against the number of incorrect ‘match’ decisions to non-matching trials. Two performance measures are obtained. The discrimination measure (d’) represents the perceptual capacity to differentiate between matching trials and non-matching trials and is obtained by calculating the difference between the standardised scores for hits (correctly saying match to a ‘matching’ trial) and false alarms (incorrectly saying match to a ‘non-matching trial). Consequently a high value represents good performance. In contrast, the response bias measure (C) represents the decision criterion used to say ‘match’ or ‘not a match’ across all trials and thus can reveal a tendency to report one outcome more than the other. It is obtained by calculating the midpoint between the standardised hit rate and standardised false alarm rate. Consequently a value of zero represents unbiased performance whilst positive values reflect a more stringent criterion in favour of ‘not a match’ decisions, and negative values reflect a more relaxed criterion in favour of ‘match’ decisions.

For the present study, however, only response bias may be of interest. This was because there was a clear *a-priori* interest in examining the effect of cognitive information on performance, but its definition as consistent or misleading depended on the ground truth of the trial (matching, non-matching). As such, collapsing performance across matching and non-matching trials was not useful. With this in mind, analyses are reported here on overall response bias, before examining accuracy of performance to explore the effects of time pressure, trial type, and contextual information.

*Bias (C)*

In order to determine whether participants displayed any bias in responding before any contextual information was introduced, a one-sample *t-*test was used to compare the overall response bias in the control condition to zero. This revealed a significant liberal bias to say ‘matching’ too often both when performance was taken as a whole (*t*(47) = -4.39, *p* < .001), and when performance was examined separately under low time pressure (*t*(23) = -2.5, *p* = .02) and under high time pressure (*t*(23) = -3.7, *p* = .001). This was perhaps not surprising as it was understandably difficult for the trained student participants to find differences between fingerprints especially when given only a brief period of time to examine the prints. An independent samples *t*-test suggested, however, that the increase in time pressure did not create a significantly greater level of response bias (*t*(46) = 1.27, *ns*).

*Accuracy*

Table 1 summarises the accuracy of performance in each of the experimental conditions, calculated with reference to ground truth. From this, it was clear that averaged levels of performance were neither at floor nor at ceiling. However, the data did reveal notable variation in performance both across participants and across items even before any contextual information was introduced. In fact, performance varied between 71% and 100% across participants and between 56% and 100% across items under a low time pressure, and it varied between 54% and 88% across participants and between 38% and 100% across items when the time pressure was increased.

(Please insert Table 1 about here)

Initial analysis took the form of a 3-way Analysis of Variance (ANOVA) to explore the effects of contextual information (DNA match, DNA unclear, DNA no-match), trial type (‘matching’, ‘non-matching’) and time pressure (high, low). Results are reported with analysis conducted both by-participants and by-items given the variation noted above.

*By-Participants:* When analysis was conducted by-participants, this revealed a main effect of trial type (*F*(1, 92) = 22.28, *p* < .001, partial η2 = .31) with performance being better on ‘matching’ than ‘non-matching’ trials overall. This reflected the response bias noted above, and the difficulty in spotting points of non-correspondence in non-matching fingerprint pairs. There was also a main effect of time pressure (*F*(1, 46) = 18.90, *p* < .001, partial η2 = .29) with performance being better when participants had unlimited inspection time compared to a time pressure. This confirmed the effectiveness of our manipulation. There was no overall effect of contextual information (*F*(2, 92) = 2.58, *p* > .05) as this necessarily interacted with the nature of the trial (*F*(2, 92) = 12.16, *p* < .001, partial η2  = .21). Importantly, the three-way interaction between all factors revealed a marginal effect suggesting that the anticipated interaction between trial type and contextual information may alter as the time pressure was increased (*F*(2, 92) = 2.77, *p* = .068, partial η2  = .06).

*By-Items:* When the same 3-way ANOVA was conducted by-items, the results reported above were replicated in all respects except that the three way interaction between all factors now reached significance (*F*(2, 70) = 5.10, *p* = .009, partial η2 = .13). Given this, four separate one-way ANOVAs were conducted to examine the effect of contextual information on ‘matching’ trials and on ‘non-matching’ trials when under low time pressure and when time pressure was increased. This approach enabled a targeted exploration of the *a-priori* expectations in each case and allowed exploration of the marginal and the significant 3-way interactions noted above. Cognitive bias would be demonstrated if the decisions were affected by the nature of the contextual information accompanying the fingerprints. Given the complexity of the results, findings are summarised in Table 2.

(Please insert Table 2 about here)

*Matching Trials*

*No time pressure:* A one-way ANOVA was conducted to examine the effect of contextual information on performance for matching trials without time pressure. For these trials, a DNA-match bias was consistent and a DNA-no match bias was misleading. Cognitive bias would be demonstrated if performance was better in the DNA-match trials and worse in the DNA-no match trials relative to the unbiased trials. The analysis revealed no significant effect of contextual information when performance was examined by participants (*F*(2, 46) = 1.69, *ns*) or by items (*F*(2, 70) = 1.92, *ns*). This most likely reflected the fact that performance was good for matching trials without a time pressure, and was not improved when consistent contextual information was provided. Moreover, the presence of the response bias maintained a high level of ‘matching’ responses even when misleading contextual information was presented. As such, no demonstration of cognitive bias emerged in this condition.

*Time pressure:* The above analysis was repeated to examine the effect of contextual information on performance for matching trials when the time pressure was increased. This revealed a significant effect of contextual information, both when performance was examined by participants (*F*(2, 46) =4.35, *p* = .019, partial η2 = .16) and when examined by items (*F*(2, 70) = 4.78, *p* = .011, partial η2 = .12). Pair-wise contrasts confirmed that performance was worse in both cases when misleading contextual information was presented compared to the control condition (by participants: *t*(23) = 2.29, *p* = .032; by items: *t*(35) = 2.16, *p* = .038). As such, cognitive bias was demonstrated, and performance was impaired, when contextual information was misleading. In contrast, the pair-wise contrasts revealed no difference in performance when consistent contextual information was presented compared to the control condition (by participants: *t*(23) < 1, *ns*; by items: *t*(35) < 1, *ns*) suggesting that performance was already high in the control condition and was not improved by a consistent suggestion.

*Non-Matching Trials*

*No time pressure:* The above analyses were repeated to examine the effect of contextual information on performance for non-matching trials. For these trials, a DNA-match bias was misleading and a DNA-no match bias was consistent. Cognitive bias would be demonstrated if performance was better in the DNA-no match trials and worse in the DNA- match trials relative to the unbiased trials. Without any time pressure, this analysis revealed a significant main effect of contextual information when performance was examined by participants (*F*(2, 46) = 4.78, *p* = .013, partial η2 = .17) and by items (*F*(2, 70) = 2.76, *p*(1-tailed) = .035, partial η2 =.07). As above, pairwise contrasts revealed that performance was worse in both cases when the misleading contextual information was presented compared to the control condition (by participants: *t*(23) = 2.73, *p* = .012; by items: *t*(35) = 2.68, *p* = .011). This again means that cognitive bias was demonstrated, and performance was impaired, when contextual information was misleading. However, again, the pair-wise contrasts revealed no benefit from presentation of the consistent contextual information compared to the control condition (by participants: *t*(23) < 1, *ns*; by items: *t*(35) < 1, *ns*). These results suggested that with performance being a little lower on the non-matching trials than on the matching trials above, the contextual information had an opportunity to exert an influence despite only a low level of time pressure.

*Time pressure:* Finally, the above analyses were repeated to examine the effect of contextual information on accuracy for non-matching trials when the time pressure was increased. This revealed a significant main effect of contextual information both by participants (*F*(2, 46) = 8.53, *p* = .001, partial η2 = .27) and by items (*F*(2, 70) = 9.75, *p* < .001, partial η2 = .22). However, in contrast to the previous effects which were carried by the influence of misleading contextual information, the pair-wise contrasts suggested that these effects reflected an improvement in performance when consistent contextual information was presented (by participants: *t*(23) = 4.55, *p* < .001; by items: *t*(35) = 3.51, *p* = .001). Indeed, performance did not differ from the control condition when misleading information was presented (by participants: *t*(23) < 1, *ns*; by items: *t*(35) < 1, *ns*). As such, the present results demonstrated cognitive bias but this time through an improvement in performance when contextual information was consistent. In reflecting on this change in characterisation of the cognitive bias, it is worth noting that performance in the unbiased (control) condition was above chance but was relatively poor when responding to non-matching trials under increased time pressure. Indeed, this condition attracted the weakest level of performance across the entire experiment. Consequently, performance could be improved by the consistent contextual information without hitting ceiling levels, whereas it could not be impaired substantially by the misleading contextual information without approaching floor level performance.

o

*Discussion*

The work presented here has examined the effect of contextual information on decision making when trained participants evaluated a pair of fingerprints. The objective was to determine whether the concerns over cognitive bias were replicated when a larger sample of participants, and trials, was used, and when task demands were manipulated in a measurable way. It was anticipated that performance would be affected by the presentation of contextual information such that, relative to an unbiased or control condition, participants would be swayed towards a ‘match’ decision when the DNA evidence indicated a match between suspect and perpetrator. Similarly, it was anticipated that they would be swayed towards a ‘not a match’ decision when the DNA evidence indicated no match between suspect and perpetrator. The accuracy or otherwise of these decisions necessarily depended on the ground truth of each fingerprint pair, with a bias to say ‘match’ being helpful in the case of a truly matching pair of prints, but being unhelpful in the case of a truly non-matching pair of prints.

Examination of the data revealed support for all predictions providing a clear demonstration of cognitive bias when participants were aware of accompanying DNA results. Even when there was no time pressure, fingerprint decisions were swayed in the presence of misleading DNA evidence. However, when a time pressure was introduced, cognitive bias was more apparent, with decisions being swayed both by misleading information (more errors on matching trials), and by consistent information (more correct decisions on non-matching trials), relative to the unbiased condition. As such, the results add to previous demonstrations of cognitive bias in a fingerprint matching task, with bias introduced here through the provision of DNA test results.

Three particular aspects of the current findings warrant mention: First, the current experiment enabled an exploration of cognitive bias not through manipulation of crime details such as the type, severity or emotionality of the crime (Dror et al., 2005; Earwaker et al., 2015; Hall & Player, 2008; Searston et al., 2015), nor through manipulation of the suspect details such as their confession or alibi (Dror & Charlton, 2006). Instead, the present manipulation of bias centred on the provision of forensic details through provision of DNA test results. This is an important finding given that DNA evidence is often hailed as the gold standard of forensic science (though see Dror & Hampikian, 2011) and as such, may be weighed more heavily by analysts compared to other contextual details in the case. The demonstration here of bias in fingerprint decisions supported our *a-priori* predictions, and complemented previous demonstrations in the field of forensic anthropology in which DNA information biased expert decisions (Nakhaeizadeh et al., 2014). However, the current study provides the first demonstration in the domain of fingerprint analysis that knowledge of DNA results can introduce cognitive bias.

Second, the current experiment enabled a thorough exploration of bias in both possible directions. More specifically, the DNA-match information encouraged participants towards a fingerprint individualisation (match) whereas the DNA-no-match information encouraged them towards a fingerprint exclusion (no-match). When performance was evaluated relative to the unbiased control condition, the results suggested that participants could be biased in both directions. As such, these results confirmed the findings of Dror and Charlton (2006) who showed that the performance of experts could be affected in both directions through introduction of a confession or an alibi context. Consequently, the present results extended the evidence-base showing cognitive bias regardless of the direction of the influence.

Third, performance in the current experiment was significantly worse when participants were under an increased time pressure. Indeed, the significant overall effect of time pressure, both by-participants and by-items, confirmed the effectiveness of this experimental manipulation. This was convenient for the present study given specific predictions regarding the effect of contextual information when task demands increased. Indeed, as performance declined with increasing time pressure, participants became more vulnerable to the effect of the contextual information such that misleading information lowered good performance (‘matching’ trials) and consistent information improved weaker performance (‘non-matching’ trials). These results confirmed those of Dror et al. (2005) who showed an effect of contextual information only when pairs of prints were highly similar and thus decisions were ambiguous or difficult. In this regard, it is worth noting that the clear-cut trials in Dror et al.’s (2005) study included pairs of prints which did not match in terms of basic fingerprint pattern (see Dror et al., 2005, figure 1 p. 802). Within an ecologically valid context, pairs such as these would be swiftly categorised as an exclusion (‘non-matching’ pair) and thus would not attract any detailed comparison. In contrast, the current study only utilised pairs of prints which were matched on their basic pattern so as to avoid a trivially easy task. In effect, all fingerprint pairs used within the current study would have fallen into Dror et al.’s (2005) ‘ambiguous’ group hence the strong effects of bias noted here.

In reflecting on the effect of time pressure, it is worth noting that this effect also has importance in its own right. Indeed, Jain and Feng (2011) noted the time pressure on latent fingerprint examiners given the backlog of cases, and the time-sensitive context in which decisions are often requested. Consequently, the increase in time pressure at the print inspection stage may reflect an ecologically valid risk factor. Here, it both weakened performance overall and rendered the participants more susceptible to the influence of contextual information. This being said, there would be value in replicating the present results under more ecologically valid conditions where decisions are likely to be taken under a less stringent level of time pressure than that employed here.

Before further discussion, there is value in remembering that the data described here were obtained from trained student participants rather than from fingerprint experts. This training was sufficient to enable the participants to perform at above-chance levels. However, as was clear from previous research (Stevenage & Pitfield, 2016), this level of training led to performance that was measurably below that of practicing fingerprint experts. This fact is helpful in the current study because performance levels were neither at floor levels (where they could not get significantly worse) nor at ceiling levels (where they could not get significantly better). As such, it was possible to bias performance in both directions within this experimental study through the presentation of directional contextual information.

Notwithstanding this, the level of training used here means that the results should not be taken as an indication of the performance levels of fully trained fingerprint experts in the field. Indeed, there are numerous illustrations of the very high levels of performance shown by experts compared to non-experts (Busey, Yu, Wyatte, Vanderkolk, Parada & Akavipat, 2011; Evett & Williams, 1995; Gutowski, 2006; Langenburg, 2009; Stevenage & Pitfield, 2016; Wertheim, Langenburg & Moenssens, 2006), even in challenging experimental conditions involving fingerprint inversion, sequential print presentation over a short timeframe, brief 2-second presentations (Thompson & Tangen, 2014, Experiments 1, 2 and 4), and genuine crime scene latent prints (Thompson, Tangen & McCarthy, 2013). This is not to say that fingerprint experts are immune to cognitive bias. Indeed, such effects have been shown when experts have been tested during the course of their day-to-day caseload (see Dror et al., 2011, Dror & Charlton, 2006; Dror et al., 2006; Fraser-Mackenzie et al., 2013; Langenburg, 2009; Langenburg et al., 2009). However, it may be the case that the magnitude of such effects differs from that reported here.

*‘Debiasing’ Fingerprint Analysis – A Summary of Recommendations.*

The results here are unambiguous in further demonstrating cognitive bias through the introduction of contextual information when completing a fingerprint matching task. Although we should be cautious about generalising the results of an experimental task such as this to a casework context, the question remains regarding the best way(s) to manage and minimise potential bias. In this regard, it is important to remember the insights of Kassin et al. (2013; see also Dror et al., 2013) who note that cognitive bias is unintentional and unconscious. An individual may have no awareness of the effects of contextual information, meaning that they cannot detect when their performance has been affected (see NCFS, 2015), and they cannot control against the effects by the act of sheer willpower (see Dror et al., 2013). Indeed, the only way to manage or minimise cognitive bias will be to change the way that the fingerprint analysts work.

In this regard, several recommendations have been articulated within the existing literature (see Dror, 2013; 2015; Kassin et al., 2013; Forensic Science Regulator, 2015; Reese, 2012) with the aim of protecting fingerprint evidence from challenges in court. Particularly valuable has been the approach of Dror (2016) who has outlined a hierarchy of expert performance (HEP) detailing 8 levels at which expert performance (in any domain) may be assessed and vulnerabilities exposed. This makes the important point that whilst performance may vary within a person (across time) or between different people, and whilst it may vary with or without the influence of bias from contextual information, a distinction should be drawn between the *observation* of the evidence, and the *conclusions* drawn from those observations. The present data provided demonstration of bias at the level of the conclusions drawn, revealing between-participant differences when making decisions to the same fingerprints as a consequence of the contextual information presented (Dror’s level 8 within the HEP framework).

In light of Dror’s (2016) HEP framework, one way forward is to examine vulnerabilities in fingerprint analysis and decision making at each level of his framework. This dovetails well with an examination the vulnerabilities and the potential solutions at each stage of the fingerprint analysis task. Consequently, the remainder of this discussion summarises the literature pertaining to bias at each stage of the fingerprint matching task, and provides suggestions to overcome these biases in the hope that this may guide a revision to the workflow (see Table 3).

(Please insert Table 3 about here)

*Analysis:* At the *Analysis* stage of the process, the concerns highlighted thus far were that contextual information through knowledge of the decisions of others (Fraser-Mackenzie et al., 2013), or through knowledge of the crime type (Earwaker et al., 2015) can affect the decision of ‘suitability’ of a print for further analysis. Following this, Dror et al., (2011) demonstrated that the availability of a comparison or reference print would affect the determination of *minutiae* within the latent print. Both represent demonstrations of cognitive bias in that contextual information irrelevant to the print itself has nevertheless affected the judgements made about the print. Both effects could, however, be minimised if the latent print was analysed in isolation of any reference prints, and if the analyst was ‘blind’ to the decisions of others. Dror (2015) notes the importance of these measures as ways to minimise ‘cognitive contamination’ and enable the analysts to better fulfil their roles as independent scientists.

When considering DNA analysis, Krane, Ford, Gilder, Inman, Jamieson, Koppl, Kornfield, Risinger, Rudin, Taylor and Thompson (2008) advocated the process of sequential unmasking in which information is revealed in sequence, and only when necessary, so as not to contaminate earlier decision stages. When applied to fingerprint analysis, this would suggest that the reference print should be presented only after the latent print has been examined and minutiae have been contemporaneously documented. However, Krane et al. (ibid) were mindful of the value of allowing the analyst to revisit their observations and make documented revisions. The advantage of such a process is that the analyst may be able to correct an oversight or error when the reference print becomes available. The danger, however, is that the subsequent decision-making may become directed more by the evidence in the reference print than by the evidence in the latent print (or is driven more by the suspect rather than the perpetrator). Given this, Dror, Thompson, Meissner, and colleagues (Dror et al., 2015) suggest the value of a balance point which provides some constraints around the revision of the latent analysis. They suggest allowing additions but not deletions of observations, or allowing changes only to low-confidence observations but not to high confidence observations. Either way, it is clear that the analysis stage of the fingerprint matching task should be conducted ‘blind’ to the decisions of others, and should be based on the latent print alone, with the potential for recorded revisions when a reference print is subsequently provided (see recommendations within the UK Forensic Science Regulator guidance).

All such revisions to the process for fingerprint analysis are designed to reduce the risk of unreliability, or cognitive bias, at Dror’s (2016) ‘observation’ stages (levels 1-4 of the HEP framework). The result may be greater consistency of observations surrounding ‘suitability’ decisions and the mark-up of minutiae, within and across analysts, and may minimise against cognitive bias through protecting against the introduction of contextual case details (though see Dror, in press, for sources of contextual bias that may continue to exist due to mere human nature, or training and operational settings).

*Comparison and Evaluation:* At these stages of the fingerprint matching task, several concerns have been highlighted. The current paper has provided one demonstration of cognitive bias in that the decisions of participants were affected by accompanying contextual information regarding the results of DNA tests related to the case. However, the literature has also demonstrated cognitive bias following the presentation of contextual information about the crime type/severity/emotionality (Dror et al., 2005; Hall & Player, 2008; Searston et al., 2015) and following information about the suspect (Dror & Charlton, 2006; Smalarz, et al., 2016). All such effects may be minimised if the analyst was ‘blind’ to these task-irrelevant factors through being shielded from case-related details (Risinger, Saks, Thompson, & Rosenthal, 2002). However, Searston et al. (2015) caution against blindly removing case-related details and assuming accuracy to improve across the board: the incidence of false alarms may be reduced, but the incidence of genuine matches may be reduced also.

Several researchers have highlighted the value of a case manager who acts independently of the analyst in a case but serves the critical role of interacting between parties involved in an investigation. Such a role necessarily protects the analyst from the irrelevant contextual information that may otherwise bias performance. However, it preserves the critical function of communication and liaison between departments involved in an active case (Found & Ganas, 2013; Dror, 2015; NCFS, 2015; Stoel, Berger, Kerkhoff, Mattijssen, & Dror, 2014; Thompson, 2011). The role of case manager and analyst may rotate across cases such that an individual may be a case manager for one case but an analyst for another. This would provide diversity to an individual’s job, and breadth to their skills set (NCFS, 2015). Indeed, this may be important in overcoming the sense of boredom or tedium as reported by laboratory managers (Butt, 2013). Similarly, computerised Laboratory Information Management Systems (LIMS) should be capable of ensuring that analysts are not exposed to task-irrelevant information too early in the process workflow.

Both approaches are sound in principle but depend upon a clear definition and identification of what is relevant (and what is not) at each stage of the task (see appendix to NCFS, 2015; Reese, 2012). In the case of fingerprint analysis, the only relevant details beyond the prints themselves may relate to the surface that the latent print was collected from, and the method of obtaining the suspect print, as both may introduce distortion and thus may be important to the comparison process. All other case details are most likely irrelevant to the task and should be withheld from the analyst until their task is completed. This said, the qualitative work described by Charlton, Fraser-Mackenzie, and Dror (2010) highlighted the motivational factors described by analysts, and included the fact that analysts reported a sense of satisfaction and joy when gaining closure on a high profile or serious case. In this regard, there is value in revealing case details to the analyst at the conclusion of their task in order to maintain motivation and job-satisfaction in what is a difficult and highly pressured role.

The *Comparison and Evaluation* stage of the fingerprint matching process has also been seen to suffer from an expectancy bias in which an analyst may conclude a ‘match’ because of the expectation that the latent print will belong to the suspect. This may arise if the analyst presumes that the police have the right suspect (a prosecutorial bias) or if they know that the suspect has a prior history, or gang affiliations (see Saks et al., 2003). Kassin et al. (2013) and Wells Wilford, and Smalarz (2013) highlighted the value of an alternative to the latent-suspect comparison as a way of addressing this expectancy bias. They described the presentation of an evidence lineup or ‘filler control method’ consisting of the suspect and five similar foils (see also Miller, 1987 for a test of evidence lineups in hair analysis, Saks, et al., 2003), with foils potentially being drawn from the similar matches produced in an AFIS search (Reese, 2012). The analyst’s task then is akin to that of an eyewitness when faced with a line-up in which the suspect is presented alongside a number of foils. In this way, the expectancy bias to say ‘match’ can be minimised as the analyst would not know which reference sample belongs to the suspect. Similarly, the influence of some irrelevant contextual details, such as suspect criminal history or confession/alibi details, can be minimised as the analyst would not know which print they related to. Finally, such a method is capable of revealing false positive errors relating to the selection of an innocent foil, and thus of providing feedback to analysts, although this method alone would not guard against the erroneous selection of an innocent suspect.

Finally, expectancy bias at this stage may also be created inadvertently through the introduction of technologies designed to ease the analyst’s task. For example, a system such as AFIS is capable of searching huge databases to retrieve close matches to a latent print. The system then outputs a list of reference prints for an analyst to consider, and these are ranked in terms of their similarity to the latent print. The notable advantage of such a system is the evident efficiency in identifying a small set of persons of interest. However, the danger is that the analyst may be primed to assume that a reference print higher up the ranked list is more likely to represent a match to the latent print compared to a reference print lower down the list. This is an example of a base-rate expectation and it rests on learned patterns from past experience. The results of Dror, Wertheim, Fraser-Mackenzie, and Walajtys (2012) confirmed this concern. Across 23 experts, the rate of false identifications was higher for reference prints higher in the AFIS list, even when the true perpetrator’s print was present lower in the list (see also Dror, 2013; 2015; Dror & Mnookin, 2010). The solution to this problem is, however, simple and involves either the randomisation of the AFIS-generated reference prints (Forensic Science Regulator, 2015), or movement of the top print further down the AFIS list (Dror, in press), in order to break the base-rate expectation. Information relating to relative similarity of each print to the latent will be lost in so doing, but so will the potential source of bias.

*Verification:* Cognitive bias has also been highlighted at the verification stage during which a second analyst checks the decision of the first analyst. The bias noted at this stage largely arises due to the fact that the second analyst is often aware of the initial decision and may then seek evidence to confirm that decision. However, there may also be an expectancy bias created if the analyst’s organisation commonly only puts ‘individualisation’ decisions forward for verification. In order to minimise these biases, the verification analyst should again be ‘blind’ to the initial decision (Campbell, 2011; Forensic Science Regulator, 2015). Ideally, they should also be blind to the identity of the initial analyst in order to avoid interpersonal issues, judgements or conflicts (Dror, 2015).

These measures may be relatively easy to implement. However, where a laboratory operates a practice of verifying ‘individualisations’ only, a strong base-rate expectancy may be created in the mind of the second analyst leading them to expect a match. Two revisions to the procedure may be envisaged to address this: The first involves the verification of all decisions regardless of the original outcome (see FBI revisions to protocols, Ruth, 2012). Alternatively, if financial and time constraints preclude this, the verification stream should include several similar non-matching fingerprint pairs to counter the expectation that verification tasks are likely to involve ‘matches’. Finally, Reese (2012) reviews the recommendation of Koppl who suggests value in ‘duplicate testing’ of a small portion of a caseload by competing laboratories within a jurisdiction, with each laboratory verifying the decisions and procedures of the other, with appropriate incentives. Whilst this may have benefit in detecting errors through waning cognitive effort, it will not remove cognitive bias and thus is not sufficient on its own. Moreover, all recommendations necessarily increase the workload of analysts, though to differing degrees. As such, Charlton (2013) and Butt (2013) very sensibly note a need for a dialogue concerning what is desirable and what is practical given fiscal limits.

The set of revisions described to the comparison, evaluation and verification stages of fingerprint analysis, are designed to reduce the risk of unreliability, or cognitive bias, at Dror’s (2016) ‘conclusion’ stages (levels 5-8 of the HEP framework). The result may be greater consistency of decision-making surrounding match, no-match, and inconclusive decisions, within and across analysts, and may again minimise cognitive bias through base rate expectations and contextual case details.

In all, the cumulative demonstration of risks to decision-making across a number of research laboratories has encouraged an urgent and systematic exploration of the stages of the vulnerabilities of the fingerprint analysis task. This has enabled a far greater appreciation of the frailties of the human expert. However, on a more positive note, it has supported a set of recommendations summarised here to provide a careful revision to procedures.

*Factors beyond the Fingerprint Analyst*

It is worthwhile noting that the extent to which these recommendations may be successful in mitigating risk is likely to be affected by factors that go beyond the task of the fingerprint analyst themselves. For instance, Reese (2012) notes that the incidence of cognitive bias is more prevalent when the task is more difficult. Consequently, she discusses the use of a different procedure, or the allocation of more time, to decisions that are deemed to be particularly difficult. This may result in a difficult fingerprint comparison being assigned a probability (rather than a certainty) of representing a match or an exclusion, with the consequence that such a judgement is used as investigative evidence only rather than being presented as legal evidence in court. Of course, there are issues over how to determine ‘difficulty’ and there may also be some resistance to a determination of difficulty within a profession which is used to expressing certainties even in taxing cases. Consequently, these suggestions are ones that the forensic community needs to weigh when considering debiasing solutions.

On a related note, the current practice for analysts to give one of three categorical decisions (100% certainty of an identification; 100% certainty of an exclusion, or an inconclusive decision) runs counter to the practice in other forensic domains. For example, in the DNA domain, probabilistic judgements are provided detailing the probability of a sample coming from a particular suspect. The concern with the categorical approach is that an individual can very often express more confidence in a decision than is warranted, with the consequence that categorical judgements fall foul of an overconfidence bias (see Fischhoff, Slovic & Lichtenstein, 1977). This mismatch between categorical decision making in the fingerprint domain, and probabilistic decision making in other forensic domains has led to calls for probabilistic decision making to be applied in the area of fingerprint analysis too (see Champod, 2011; Campbell, 2011; Mnookin, 2008; Neumann & Champkin, 2012). However, again a resolution of this issue is one that extends beyond the fingerprint task itself, and it will require careful consideration by the forensic community.

Finally, Dror (2015) notes that cognitive bias may arise through adversarial allegiance, in which decision making is affected according to whether an expert is retained by the defence or the prosecution (Murrie, Boccaccini, Guarnera & Rufino, 2013, cited in Dror, 2015). This, of course, exists as a consequence of the court adversarial system, and is a source of bias that applies beyond the domain of fingerprint analysis. Nevertheless, it can have an unconscious impact on the extent to which an expert in any forensic domain can act as an independent scientist. This is not at all simple to address beyond the stipulation of an overriding responsibility to the court, and a code to act with honesty, integrity and impartiality (Forensic Science Regulator, 2015). However, given that the bias may operate at an unconscious rather than a conscious level, mere reinforcement of these standards may not be effective as a debiasing measure. This said, it would be interesting to determine the extent of this adversarial allegiance effect relative to any bias that may result from mere court appearance (such as under the inquisitorial court system as operates in parts of Europe).

*Summary and Conclusions*

The purpose of the present paper was to determine whether participant decisions on a fingerprint matching task would be affected by the presentation of DNA evidence related to the case. Trained but non-expert participants completed a large number of fingerprint matching trials in which ground truth was known. Critically, however, some trials were accompanied by DNA evidence designed to suggest that the suspect and perpetrator matched, whilst other trials were accompanied by DNA evidence designed to suggest the opposite. Performance in each condition was compared to that in an unbiased control condition where the DNA evidence was inconclusive. The results suggested that, even when there was no time pressure, performance on non-matching trials was impaired if misleading DNA evidence was presented. However, when a time pressure was introduced, then misleading information impaired performance (matching trials) and consistent DNA evidence improved performance (non-matching trials) relative to the unbiased condition. These results are the first to demonstrate the impact of DNA results on fingerprint decision making. However, the pattern of performance echoes previous demonstrations of cognitive bias in the fingerprint analysis task.

The present discussion has summarised existing wisdom regarding a set of solutions which may minimise cognitive bias through a revision to the workflow. These represent sensible counter-measures (see Cole, 2013) and to this end their implementation now sits with the forensic science community who must balance their desirability against any practicality concerns. Whilst not necessarily widespread, cognitive bias is real. Given the fact that every case counts when life and liberty are concerned, there is a responsibility to minimise errors wherever possible. The hope here is that by contributing to the literature, this research effort will improve rather than threaten the credibility of the fingerprint evidence (Koehler, 2008) by making the unknown known, and by encouraging a practice which stands up to scrutiny rather than blind or biased acceptance (Triplett, 2013).

*References*

Ashbaugh, D. (1999). Quantitative-Qualitative friction ridge analysis: An introduction to basic and advanced ridgeology. Boca Raton, Fl: CRC Press.

Busey, T.A., Yu, C., Wyatte, D., Vanderkolk, J., Parada, F., & Akavipat, R. (2011). Consistency and variability among latent print examiners as revealed by eye tracking methodologies. *Journal of Forensic Identification, 61(1)*, 60-92.

Butt, L. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions – Commentary by a forensic examiner. *Journal of Applied Research in Memory and Cognition, 2*, 59-60. doi:10.1016/j.jarmac.2013.01.012

Campbell, A. (2011). The fingerprint inquiry report. *Edinburgh, Scotland: APS Group Scotland*, *790*.

Champod, C. (2011). Fingerprint examination: Towards more transparency. *Law, Probability and Risk. 7,* 165-189.

Charlton, D. (2013). Standards to avoid bias in fingerprint examination? Are such standards doomed to be based on fiscal expediency? *Journal of Applied Research in Memory and Cognition*, *2*(1), 71-72. doi:10.1016/j.jarmac.2013.01.009

Charlton, D., Fraser-Mackenzie, P. A., & Dror, I.E. (2010). Emotional experiences and motivating factors associated with fingerprint analysis. *Journal of Forensic Sciences, 55*(2),385-393. doi:10.1111/j.1556-4029.2009.01295.x

Charman, S.D., Gregory, A.H., & Carlucci, M. (2009). Exploring the diagnostic utility of facial composites: Beliefs of guilt can bias perceived similarity between composite and suspect. *Journal of Experimental Psychology: Applied, 15*, 76-90. doi: 10.1037/a0014682

Cole, S.A. (2005). More than zero: Accounting for error in latent fingerprint identification. *Journal of Criminal Law and Criminology,* 95, 985-1078. doi: 0091-4169/05/9503-0985

Cole, S.A. (2013). Implementing counter-measures against confirmation bias in forensic science. *Journal of Applied Research in Memory and Cognition, 2*, 61-62. doi: 10.1016/j.jarmac.2013.01.011

Data in Brief (submitted). Training and performance measures for novices to the area of fingerprint analysis. *Forensic Science International: Data in Brief*.

Dror, I.E. (in press). Human expert performance in forensic decision making: Seven different sources of bias. *Australian Journal of Forensic Science*. http://dx.doi.org/10.1080/00450618.2017.1281348

Dror, I.E., (2016). A hierarchy of expert performance. *Journal of Applied Research in Memory and Cognition, 5*, 121-127.

Dror, I.E., (2015). Cognitive neuroscience in forensic science: Understanding and utilizing the human element. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences, 370,* 20140255. doi: 10.1098/rstb.2014.0255

Dror, I.E., (2013). Practical solutions to cognitive and human factor challenges in forensic science. *Forensic Science Policy & Management, 4(3-4)*, 1-9. doi: 10.1080/19409044.2014.901437

Dror, I.E., (2012). Cognitive bias in forensic science. In *The 2012 Yearbook of Science & Technology.* New York: McGraw-Hill, pp 43-45.

Dror, I.E., (2009). On proper research and understanding of the interplay between bias and decision outcomes. *Forensic Science International, 191*, e17-18. doi: 10.1016/j.forsciint.2009.03.012

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, *56*(4), 600-616.

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, *156*(1), 74-78. doi:10.1016/j.forsciint.2005.10.017

Dror, I. E., Champod, C., Langenburg, G., Charlton, D., Hunt, H., & Rosenthal, R. (2011). Cognitive issues in fingerprint analysis: inter-and intra-expert consistency and the effect of a ‘target’comparison. *Forensic Science International*, *208*(1), 10-17. doi: 10.1016/j.forsciint.2010.10.013

Dror, I.E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science and Justice, 51*, 204-208. doi: 10.1016/j.scijus.2011.08.004

Dror, I.E., Kassin, S.M., & Kukucka, J. (2013). New application of psychology to law: Improving forensic evidence and expert witness contributions. *Journal of Applied Research in Memory and Cognition, 2*, 78-81. doi: 10.1016/j.jarmac.2013.02.003

Dror, I.E., & Mnookin, J. (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 and Risk, 9*, 47-67. doi: 10.1093/lpr/mpg031

Dror, I. E., Péron, A. E., Hind, S. L., & Charlton, D. (2005). When emotions get the better of us: the effect of contextual top‐down processing on matching fingerprints. *Applied Cognitive Psychology*, *19*(6), 799-809. doi:10.1002/acp.1130

Dror, I.E., Thompson, W.C., Meissner, C.A., Kornfield, I., Krane, D., Saks, M., & Risinger, J. (2015). Context management toolbox: A linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. *Journal of Forensic Sciences, 60(4)*, 1111-1112.doi: 10.1111/1556-4029.12805

Dror, I.E., Wertheim, K., Fraser-Mackenzie, P., & Walajtys, J. (2012). The impact of human-technology cooperation and distributed cognition in forensic science: Biasing effects of AFIS contextual information on human experts. *Journal of Forensic Sciences, 57(2)*, 343-352. doi: 10.1111/j.1556-4029.2011.02013.x

Earwaker, H., Morgan, R.M., Harris, A.J.L., & Hall, L.J. (2015). Fingermark submission decision-making within a UK fingerprint laboratory: Do experts get the marks that they need? *Science and Justice, 55*, 239-247.

Evett, I., & Williams, R. (1995). A review of the sixteen points fingerprint standard in England and Wales. In *Proceedings of the International Symposium on Fingerprint Detection and Identification*. J. Almog, & E. Springer, E. (Eds.). Ne’urim, Israel, pp 287 – 304.

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

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3(4)*, 552-564.

Forensic Science Regulator (2015). Cognitive bias effects relevant to forensic science examinations: Guidance. FSR-G-217 (Issue 1) https://www.gov.uk/government/uploads/system/uploads/attachment\_data/file/510147/217\_FSR-G-217\_Cognitive\_bias\_appendix.pdf

Found, B., & Ganas, J. (2013). The management of domain irrelevant context information in forensic handwriting examination casework. *Science & Justice, 53(2)*, 154-158. doi: 10.1016/j.scijus.2012.10.004

Fraser-Mackenzie, P.A., Dror, I.E., & Wertheim, IK. (2013). Cognitive and contextual influences in determination of latent fingerprint suitability for identification judgments. *Science and Justice, 53(3)*, 144-153. doi: 10.1016/j/scijus.2012.12.002

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D.J. Koehler & N. Harvey (Eds). *Blackwell Handbook of Judgment and Decision Making.*

Green, D.M., & Swets, J.A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Gutowski, S. (2006). Error rates in fingerprint examination: The view in 2006. *The Forensic Bulletin, Autumn, 2006,* 18-19.

Hall, L. J., & Player, E. (2008). Will the introduction of an emotional context affect fingerprint analysis and decision-making? *Forensic Science International*, *181*(1), 36-39. doi:10.1016/j.forsciint.2008.08.008

Jain, A.K., & Feng, J. (2011). Latent fingerprint matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(1)*, 88-100. doi: 10.1109/TPAMI.2010.59

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, *2*(1), 42-52. doi:10.1016/j.jarmac.2013.01.001

Koehler, J.J. (2008). Fingerprint error rates and proficiency tests: What they are and why they matter. *Hastings Law Journal, 59*, 1077-1100.

Krane, D.E., Ford, SA., Gilder, J., Inman, K., Jamieson, A., Koppl, R., Kornfield, I.L., Risinger, D.M., Rudin, N., Taylor, M.S., & Thompson, W.C. (2008). Sequential unmasking: A means of minimizing observer effects in forensic DNA interpretation. *Journal of Forensic Science, 53(4)*, 1006-1007. doi: 10.1111/j.1556-4029.2008.00787.x

Langenburg, G., (2009). A performance study of the ACE-V Process: A pilot study to measure the accuracy, precision, reproducibility, repeatability, and bias ability of conclusions resulting from the ACE-V process. *Journal of Forensic Identification, 59(2)*, 219-257.

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 of Forensic Science, 54(3)*, 571-582. doi: 10.1111/j.1556-4029.2009.01025.x

Miller, L.S. (1987). Procedural bias in forensic science examinations of human hair. *Law and Human Behavior, 11(2),* 157-163.

Mnookin, J.L. (2008). The validity of latent fingerprint identification: Confessions of a fingerprinting moderate. *Law Probability and Risk, 7*, 127-141.

Nakhaeizadeh, S., Dror, I.E., & Morgan, R.M. (2014). Cognitive bias in forensic anthropology: Visual assessment of skeletal remains is susceptible to confirmation bias. *Science and Justice, 54*, 208-214. doi:10.1016/j.scijus.2013.11.003

National Academy of Sciences (2009). *Strengthening forensic science in the United States: A path forward.* Washington, DC. National Academies Press.

National Commission on Forensic Science (2015). Ensuring that forensic analysis is based upon task-relevant information. Available at https://www.justice.gov/ncfs/file/641676/download

Neumann, C., & Champkin, J. (2012). Fingerprints at the crime-scene: Statistically certain or probable? *Significance*: *The Royal Statistical Society, February 2012,* 21-25.

OIG. (2006). *A review of the FBI’S handling of the Brandon Mayfield case*. Retrieved from <https://oig.justice.gov/special/s0601/final.pdf>

Pankanti, S., Prabhakar, S., & Jain, A.K. (2002). On the individuality of fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(8),* 1010-1025. doi: 10.1109/TPAMI.2002.1023799

Reese, E.J. (2012). Techniques for mitigating cognitive biases in fingerprint identification. *UCLA Law Review, 59*, 1252-1290.

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 suggestions. *California Law Review, 90*, 1-55. Available at http://nrs.harvard.edu/urn-3:HUL.InstRepos:3199069

Ruth, R.M. (2012). FBI Laboratory’s reaction to the Madrid bombing errors – Reviews and Changes. Paper presented at ‘Fingerprints: A Roadmap for Reform’. University of Strathclyde, 19-21st March 2012. http://ewds.strath.ac.uk/suii/Programmes/Pastprogrammes/Programmes20112012/FingerprintsARoadmapforReform.aspx

Saks, M. J., (2009). Concerning L.J. Hall, E. Player. Will the instruction of an emotional context affect fingerprint analysis and decision-making? Letter to the Editor. *Forensic Science International, 191,* e19. doi: 10.1016/j.forsciint.2009.06.011

Saks, M.J., Risinger, D.M., Rosenthal, R., & Thompson, W.C. (2003). Context effects in forensic science: A review and application of the science of science to crime laboratory practice in the United States. *Science & Justice, 43*, 77-90. doi: 10.1016/S1355-0306(03)71747-x

Schiffer, B., & Champod, C. (2007). The potential (negative) influence of observational biases at the analysis stage of fingermark individualisation. *Forensic Science International*, *167*(2), 116-120. doi:10.1016/j.forsciint.2006.06.036

Searston, R.A., Tangen, J.M., & Eva, K.W. (2015). Putting bias into context: The role of familiarity in identification. *Law and Human Behavior, 40*, 50-64. doi: 10.1037/lhb0000154

Smalarz, L., Madon, S., Yang, Y., Guyll, M., & Buck, S. (2016). The perfect match: Do criminal stereotypes bias forensic evidence analysis? *Law and Human Behavior, 40(4)*, 420-429. doi: 10.1037/lhb0000190

Stammers, S., & Bunn, S. (2015). Unintentional bias in forensic investigation. POSTbrief, Number 15, Houses of Parliament, Parliamentary Office of Science and Technology.

Stevenage, S.V., & Pitfield, C. (2016). Fact or Friction: Examination of the transparency, reliability and sufficiency of the ACE-V method of fingerprint analysis. *Forensic Science International, 267*, 145-156. doi: [10.1016/j.forsciint.2016.08.026](http://dx.doi.org/10.1016/j.forsciint.2016.08.026)

Stoel, R.D., Berger, C.E., Kerkhoff, W., Mattijssen, E., & Dror, I. (2014). Minimizing contextual bias in forensic casework. In: Strom, K., & Hickman, M.J. (Eds.). *Forensic Science and the Administration of Justice*. pp 67-86. New York: Sage.

Thompson, W.C., (2011). What role should investigative facts play in the evaluation of scientific evidence? *Australian Journal of Forensic Science, 43(2-3)*, 123-134. doi: 10.1080/00450618.2010.541499

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)*: e114759. doi: 10.1371/journal.pone.0114759

Thompson, M.B., Tangen, J.M., & McCarthy, D.J. (2013). Human matching performance of genuine crime scene latent fingerprints. *Law and Human Behavior, 38*, 84-93. doi: 10.1037/lhb0000051

Triplett, M. (2013). Errors in forensics: Cause(s) and solutions. *Journal of Applied Research in Memory and Cognition*, *2*, 63-64. doi: 10.1016/j.jarmac.2013.01.007

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185(4157),* 1124-1131. doi: 10.1126/science.185.4157.1124

Wells, G.L., Wilford, M.M., & Smalarz, L. (2013). Forensic science testing: The forensic filler-control method for controlling contextual bias, estimating error rates, and calibrating analysts’ reports. *Journal of Applied Research in Memory and Cognition, 2*, 53-55. doi: 10.1016/j.jarmac.2013.01.004

Wertheim, K., Langenburg, G., & Moenssens, A. (2006). A report of latent print examiner accuracy during comparison training exercises. *Journal of Forensic Identification, 56(1)*, 55-93.

Table Legend

Table 1:

Mean response bias (C) on a fingerprint matching task, together with the total number of correct responses and the proportion of accurate responses across ‘matching’ and ‘non-matching’ trials as evaluated against the ground truth of each trial. Data are separated out according to the nature of the DNA information accompanying the trials, with 12 matching trials and 12 non-matching trials in each of the three DNA contexts. Data are separated further according to trial type (matching, non-matching) given that the directional influence of the accompanying DNA information depends on the trial type. A total of 24 participants took part under time pressure and no time pressure conditions. Thus, the total number of correct decisions in each condition is expressed out of 288 (12 trials x 24 participants). Standard deviation is shown brackets where relevant.

Table 2:

Summary of conditions under which cognitive bias is demonstrated in the fingerprint analysis task due to the knowledge of DNA information.

Table 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | DNA-match | DNA-unclear | DNA-no match |
| Response Bias (C) | No Time Pressure  Time Pressure | -.31 (.51)  -.31 (.54) | -.21 (.41)  -.38 (.51) | -.14 (.46)  -.14 (.40) |
| Matching Trials | No Time Pressure  Tot. No. Correct  Accuracy | 254 / 288  .88 (.09) | 258 / 288  .90 (.09) | 245 / 288  .85 (.10) |
|  | Time Pressure  Tot. No. Correct  Accuracy | 240 / 288  .83 (.11) | 237 / 288  .82 (.18) | 214 / 288  .74 (.14) |
| Non-matching Trials | No Time Pressure  Tot. No. Correct  Accuracy | 209 / 288  .73 (.17) | 233 / 288  .81 (.12) | 236 / 288  .82 (.13) |
|  | Time Pressure  Tot. No. Correct  Accuracy | 186 / 288  .65 (.21) | 185 / 288  .64 (.15) | 231 / 288  .80 (.13) |

Table 2:

|  |  |  |
| --- | --- | --- |
| MATCHING  TRIALS | DNA-match  (consistent) | DNA- no match  (misleading) |
| No Time Pressure  Analysis by participants  Analysis by items | No bias  No bias | No bias  No bias |
|  | Performance was strong for matching trials with no time pressure, and was neither helped by consistent information, nor impaired by misleading information. | |
| Time Pressure  Analysis by participants  Analysis by items | No bias  No bias | BIAS  BIAS |
|  | Performance was good for ‘matching’ trials under time pressure but was impaired by misleading information under time pressure | |
| NON-MATCHING TRIALS | DNA-match  (misleading) | DNA- no match  (consistent) |
| No Time Pressure  Analysis by participants  Analysis by items | BIAS  BIAS | No bias  No bias |
|  | Performance was weaker in non-matching trials and was impaired by misleading information even with no time pressure | |
| Time Pressure  Analysis by participants  Analysis by items | No bias  No bias | BIAS  BIAS |
|  | Performance was weakest for non-matching trials under time pressure but was improved by consistent information | |

Table 3: A summary of the effects of bias, and of mechanisms to minimise bias at each stage of the ACE-V method of fingerprint matching

|  |  |  |  |
| --- | --- | --- | --- |
| Stage | Illustration of Bias | Type of Bias | Recommendations |
| Analysis | Influence from others regarding print ‘suitability’ (Fraser-Mackenzie et al., 2013).  Influence of crime type when determining print ‘suitability’ (Earwaker et al., 2015).  Influence of reference print when detailing minutiae (Dror et al., 2011). | Confirmatory bias  Contextual bias  Contextual bias | Evaluate the print ‘blind’ to the views of others.  Reinforce importance of the same thresholds for sufficiency decisions regardless of crime type.  Evaluate the print ‘solo’, with (linear) sequential unmasking to document changes to initial observations. |
| Comparison and Evaluation | Influenced by order of prints in the AFIS-generated list (Dror et al., 2012).  Expectation that a suspect will be the perpetrator (Kassin et al., 2013; Saks et al., 2003; Wells et al. 2013).  Influenced by irrelevant case details:   * Forensic details (here). * Case details (Dror et al., 2005; Searston et al., 2015; cf Hall & Player, 2008). * Suspect details (Dror & Charlton, 2006; Smalarz et al., 2016). | Expectancy bias  Expectancy bias  Contextual bias | Randomise the AFIS list of reference prints.  Shift from a suspect-perpetrator comparison to an evidence-pack (where practicable).  Case management to enable analyst to complete their task ‘blind’ to irrelevant details. |
| Verification | Influenced by knowledge of original decision (Dror et al., 2006, Langenburg et al., 2009).  Expectation that a verification task will yield a ‘match’ decision. | Confirmatory bias  Expectancy bias | Conduct verification ‘blind’ to original decision.  Verify all decisions/Introduce non-matching pairs (where practicable). |

Figure 1:

Example stimuli used in a matching trial and a non-matching trial.

|  |  |  |
| --- | --- | --- |
| Matching Trial |  |  |
| Non-Matching Trial |  |  |

1. This manipulation was considerable, and created a task which consequently bore much less resemblance to the real task of analysts in the field. However, the adoption of this manipulation was successful in ensuring a measurable effect (performance under time pressure was substantially reduced), without performance falling to a chance level akin to guessing. As such, it provided a robust experimental test of cognitive bias under demanding conditions whilst avoiding complications due to floor effects. [↑](#footnote-ref-1)
2. Whilst the provision of information was effective in enabling participants to complete the task with a degree of accuracy, this form of training obviously omitted the lengthy training, practice and feedback elements available to experts in the field. As a consequence, the results are not intended to indicate absolute levels of performance of fingerprint experts. Nevertheless, the relative performance of the trained students here across contextual information conditions remains valuable as an experimental test of the impact of DNA results on a fingerprint matching task. [↑](#footnote-ref-2)