

## **Contextual bias and cross-contamination in the forensic sciences: the corrosive implications for investigations, plea bargains, trials and appeals**

GARY EDMOND<sup>†</sup>

*Professor, Australian Research Council (ARC) Future Fellow; and Director, Expertise, Evidence & Law Program, School of Law, The University of New South Wales, Sydney 2052; Australia and Professor (fractional), School of Law, Northumbria University, Newcastle City Campus, Ellison Pl, Newcastle upon Tyne, Tyne and Wear NE1 8ST, United Kingdom*

JASON M. TANGEN AND RACHEL A. SEARSTON

*School of Psychology, University of Queensland, Brisbane, St Lucia QLD 4072, Australia*

AND

ITIEL E. DROR

*Centre for the Forensic Sciences, University College London (UCL), Gower St, London WC1E 6BT, United Kingdom*

[Received on 20 February 2014; accepted on 31 August 2014]

Most forensic science evidence is produced in conditions that do not protect the analyst from contextual information about the case that could sway their decision-making. This article explores how these largely unrecognized threats raise real problems for the criminal justice system; from the collection and interpretation of traces to the presentation and evaluation of evidence at trial and on appeal. It explains how forensic analysts are routinely exposed to information (e.g. about the investigation or the main suspect) that is not related to their analysis, and not documented in their reports, but has been demonstrated to affect the interpretation of forensic science evidence. It also explains that not only are forensic analysts gratuitously exposed to such ‘domain-irrelevant’ information, but their own cognitively contaminated interpretations and opinions are then often unnecessarily revealed to other witnesses—both lay and expert. This back and forth can create a ‘biasing snowball effect’ where evidence is (increasingly) cross-contaminated, though represented, at trial and on appeal, as separate lines of evidence independently corroborating one another. The article explains that lawyers and courts have not recognized how contextual bias and cognitive processes may distort and undermine the probative value of expert evidence. It suggests that courts should attend to the possibility of contextual bias and cross-contamination when admitting and evaluating incriminating expert evidence.

*Keywords:* expert evidence; context effects; confirmation bias; cognitive science; human factors; expectancy effects; corroboration; suggestion; priming; proof.

### **1. Contextual bias, cross-contamination and criminal justice**

This article explores the pernicious, though largely unrecognized, influence that contextual factors and cognitive processes may exert on the production of incriminating expert evidence and its presentation

<sup>†</sup>Corresponding author. E-mail: g.edmond@unsw.edu.au

and evaluation in criminal proceedings.<sup>1</sup> Drawing on decades of research from the cognitive sciences, we explain how contemporary legal practice has been insensitive to processes that threaten to subvert expert evidence and proof. Specifically, the article explains how the manner in which much forensic science evidence is produced and presented needlessly introduces real risks of error.

Many forensic scientists are routinely exposed to information that is not relevant to their processing and interpretation of evidence. Exposure to this *domain-irrelevant* information (e.g. about the suspect, police suspicions and other aspects of the case) threatens the interpretation and value of their opinion evidence. The manner in which forensic scientists are regularly and unnecessarily exposed to domain-irrelevant information is rarely disclosed, raised or considered in plea negotiations, admissibility decision-making or when different types of evidence are combined and assessed during investigations, trials and appeals. The fact that relatively few forensic scientists are actively shielded from information with the potential to mislead is hardly ever raised by prosecutors and judges or considered by jurors. In consequence, even though incriminating expert evidence is routinely developed in conditions that are known to produce errors, it is nevertheless portrayed as independent, objective and sometimes even ‘error-free’.

Whereas forensic scientists, lawyers and judges have been slowly sensitized to the dangers of physical contamination, the dangers posed by cognitive contamination (where interpretations and judgments are swayed, often without awareness or conscious control, by contextual cues, irrelevant details of the case, prior experiences, expectations and institutional pressures) affecting the interpretation and evaluation of evidence, have not received serious consideration in trials and appeals.<sup>2</sup> The cognitive processes that stand at the centre of many forensic science techniques—such as comparing and interpreting traces and data—are not protected from the risks posed by exposure to extraneous contextual information and other factors that may contaminate an analyst’s performance. Lack of attention to cognitive processes, in conjunction with the continuing exposure of many forensic analysts to domain-irrelevant information (often in the guise of explicit suggestion), threatens the value of expert evidence and legal decision-making.

Adding to the complexity and dangers, forensic science evidence is routinely represented—in investigations, plea negotiations, trials and appeals—as independent corroboration for other strands of incriminating evidence. Claims about independence and corroboration persist even where the expert evidence may have been influenced by the other strands of evidence (and vice versa).<sup>3</sup>

The problem is not only that forensic science evidence can be biased (by what the detective tells the examiners, the context of the case, and so on), but that it can bias other lines of evidence. For example, if one piece of forensic evidence (biased or not) is known to other forensic examiners who are analyzing other forensic evidence, then their examination may be affected and biased by their knowledge of the results of the other

<sup>1</sup> Because these issues infect all jurisdictions where humans perceive information, make judgments and interpret, we have intentionally kept the article general in nature.

<sup>2</sup> We identified no sustained discussion or responses to ‘contextual bias’ or ‘cognitive bias’ in reported appellate judgments in England, Australia and Canada, though there are several passing references in *Laing & Anor v. R* [2013] EWCA Crim 1836, [49]; *Resolution Chemicals Ltd v. H Lundbeck A/S* [2013] EWHC 3160, [54], [66]; *Webber & Hatton* [2013] FamCA 15; *Fonteyn v. Candetti Constructions Pty Ltd* [2010] SAIRC 43; *R v. Wiens* 2013 BCSC 1539.

<sup>3</sup> For example, a fingerprint analyst conducting a comparison might know that the suspect, whose prints are being examined, made a confession and an eyewitness might be told that the person, who they thought looked like the offender, was ‘confirmed’ by fingerprint evidence. This kind of feedback tends to strengthen the confidence of the witnesses (especially eyewitnesses), but has no obvious correlation with accuracy.

piece of evidence (for example, a forensic examiner looking at bite marks may be influenced and biased in their examination if they know that fingerprint evidence shows the suspect is guilty).

Forensic evidence can also bias other lines of evidence. For example, eyewitnesses can be affected. . . .

When they affect and influence one another, then their value and reliability is diminished. Furthermore, because one piece of evidence influences another, then greater distortive power is gathered as more evidence is affected (and affecting) other lines of evidence, causing an increasing snowball of bias.<sup>4</sup>

The dangers posed by cross-contamination, our *snowball effect*, go in all directions—non-scientific evidence influencing forensic science, the results of forensic science analyses influencing the evidence of non-expert witnesses, as well as the results of one forensic science analysis influencing the results of another. Endeavouring to capture the seriousness of the threat to proof, Simon has described non-independent ‘corroboration’ as ‘pseudo-corroboration’.<sup>5</sup> The failure to shield forensic scientists from information that is not required for their analyses threatens the objectivity, independence, impartiality and the value of their evidence as well as the standard of criminal proof.

## 2. Contextual bias and cognitive contamination: the scientific research

This section begins with a brief overview of scientific research on contextual influences, biases and the (pliable) nature of human interpretation and decision-making, then shifts to focus on the forensic sciences.

### 2.1 *The human mind and cognitive architecture: contextual influences and biases*

Human perception and memory do not operate like a video camera that reliably captures and stores every detail of our experience. They do not provide direct (or unmediated) access to our world nor allow us to re-‘view’ perceptions at some later stage. Instead, the way we perceive the world and remember events is shaped by our experiences and beliefs as well as the context and stimuli. We use our prior knowledge and contextual cues to help sift, sort and weigh the vast amount of information delivered through our senses. Our perception is moulded with each new experience and the lens through which we see the world, adjusted. We evolved, and continue to live, in a highly complex and variable environment. In order to navigate the complexity, we chronically make inferences on the basis of simplified models and heuristics. In coping with the normal stimuli of our everyday lives, we are crucially dependent on processes and knowledge of which we are often unaware.

<sup>4</sup> DROR, I. E. (2012). Cognitive Bias in Forensic Science. *Yearbook of Science and Technology* (McGraw-Hill), pp. 43–45. See also 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, 42–52; and DROR, I. E. & STOEL, R. (2014). Cognitive Forensics: Human Cognition, Contextual Information and Bias. *Encyclopedia of Criminology and Criminal Justice*, Springer, New York, 353–363. It is not our intention to suggest that evidence, particularly opinion evidence, has a proper value, but rather to recognize an appropriate (or indicative) range of values based on formal evaluation.

<sup>5</sup> See SIMON, D. (2012). *In Doubt: The Psychology of the Criminal Justice Process*, Harvard University Press, Cambridge, MA.

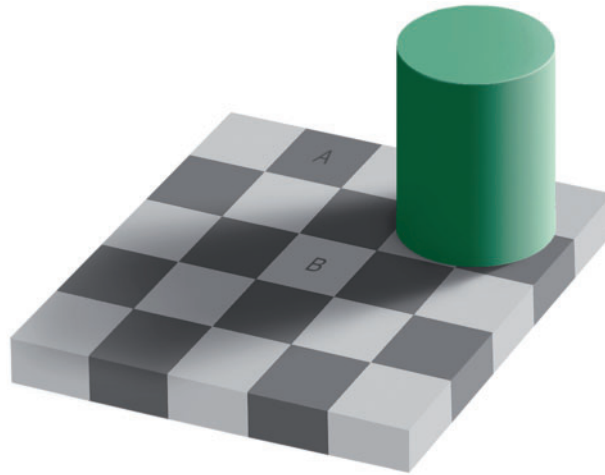


FIG. 1. The checker shadow illusion by Edward H. Adelson<sup>8</sup>.

The fact that environmental factors (or attributes of the stimulus or situation) can influence our perception and interpretation of an object or event is often referred to as a *context effect*.<sup>6</sup> Such effects are notorious and widespread.<sup>7</sup> One of the most basic exemplifications is a simultaneous contrast effect where context changes the appearance of an object (e.g., lightness, length, area, orientation, colour, etc.), but the physical properties of the object remain unchanged. For example, a grey patch will appear whiter against a dark background than against a white background. An example of a context effect is depicted in Figure 1. The squares marked A and B are the same shade of grey, yet B appears to be lighter than A. Why? Because, in a three-dimensional world, objects in shadow (e.g. B) reflect less light than objects in full illumination (e.g. A). We treat the two-dimensional image on the page ‘as if’ it comprises three-dimensional objects in the world where objects cast shadows and reflect light.<sup>8</sup>

Similar contrast effects can be observed by, for example, staring at a green patch. It will make a subsequent grey patch appear pinkish.<sup>9</sup> Staring at a face that is fat, happy, contracted, male, etc., causes successive neutral faces to appear thin, sad, expanded, female, etc.<sup>10</sup> The properties of one perceptual

<sup>6</sup> The term context effect is a general description that encompasses all aspects of the stimulus itself, the situation, and the previous experience or expectations of the subject. Others have grouped together a variety of terms to refer to this general phenomenon (e.g. observer effects, expectancy effects, cueing, priming, top-down processing, constructive perception, perceptual set, etc.). For a general review and discussion of context effects in forensic science, see 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; RISINGER, D. M., SAKS, M. J., THOMPSON, W.C. & ROSENTHAL, R. (2002). The *Daubert/Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion. *California Law Review*, **90**, 1–56.

<sup>7</sup> NICKERSON, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, **2**, 175.

<sup>8</sup> ADELSON, E. H. (1995). Checkershadow Illusion, <http://persci.mit.edu/gallery/checkershadow> [Accessed 1 October 2014].

<sup>9</sup> HEINEMANN, E.G. (1955). Simultaneous Brightness Induction as a Function of Inducing-and Test-Field Luminances. *Journal of Experimental Psychology*, **50**, 89.

<sup>10</sup> For a review of the face adaptation effect, see HILLS, P. J., HOLLAND, A. M. & LEWIS, M. B. (2010). Aftereffects for Face Attributes with Different Natural Variability: Children are More Adaptable than Adolescents. *Cognitive Development*, **25**,

stimulus change the way that we perceive or interpret subsequent stimuli. These phenomena are not limited to perceptual contexts, but also to expectations.<sup>11</sup>

Another type of context effect is *priming* (also known as cueing or suggestion), where exposure to some stimulus (such as a set of words or images or information) can influence subsequent judgments, decisions and choices. For example, if you read the word ‘eat’, and are then asked to complete the word S\_\_P, you are much more likely to fill in the blanks to read ‘soup’ than ‘soap’. On the other hand, if you are primed by the idea or the word ‘washing’ you are more likely to fill in the blanks to read ‘soap’.<sup>12</sup> People who are primed with the concept of ‘money’ (i.e., by completing a word descrambling task related to money) tend to (temporarily) behave and respond to questions in a more selfish, self-reliant and individualistic manner.<sup>13</sup> These priming effects are commonplace across experimental psychology, where activating a particular concept, attitude or goal can change the context of the situation and thereby alter how an individual responds.<sup>14</sup>

Context effects are not limited to cues in the environment. Our body and actions shape the way that we perceive and interpret the world.<sup>15</sup> Most of the time, contextual cues help us to make appropriate judgments and decisions. Our perceptual and cognitive systems go to a lot of trouble to ensure that what we see, hear and remember responds to what is usually out there.<sup>16</sup>

Research in cognitive science over the last several decades has revealed many regularities in judgment and decision-making. These are often described as *heuristics*, *biases* or *effects*, to reflect the systematic recurrence of these phenomena across individuals and circumstances, and as *fallacies* or *errors* when they lead to error.<sup>17</sup> Research demonstrates that these regularities in judgment and decision-making are usually quite useful. They can, however, be detrimental, encouraging short cuts and mistakes. Indeed, even though the term ‘cognitive bias’ is neutral, simply capturing a systematic deviation—true for virtually every judgment we make—the term is typically used pejoratively, to refer to the errors that occasionally result.

278–289; and for a related interesting face contrast effect, see TANGEN, J. M., MURPHY, S. C. & THOMPSON, M. B. (2011). Flashed Face Distortion Effect: Grotesque Faces from Relative Spaces. *Perception*, **40**, 628–630.

<sup>11</sup> BRESSAN, P. & DAL MARTELLO, M. F. (2002). *Talis Pater, Talis Filius*: Perceived Resemblance and the Belief in Genetic Relatedness. *Psychological Science*, **13**, 213–218.

<sup>12</sup> See KAHNEMAN, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York for a more detailed description of priming and related experiments.

<sup>13</sup> VOHS, K. D., MEAD, N. L. & GOODE, M. R. (2006). The Psychological Consequences of Money. *Science*, **314**, 1154–1156.

<sup>14</sup> DIJKSTERHUIS, A., SMITH, P. K., VAN BAAREN, R. B. & WIGBOLDUS, D.H. (2005). The Unconscious Consumer: Effects of Environment on Consumer Behavior. *Journal of Consumer Psychology*, **15**, 193–202.

<sup>15</sup> KILLEEN, P. R. & GLENBERG, A. M. (2010). Resituating Cognition. *Comparative Cognition & Behavior Reviews*, **5**, 59–77; STRACK, F., MARTIN, L. L. & STEPPER, S. (1988). Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis. *Journal of Personality and Social Psychology*, **54**, 768; see GLENBERG, A. M., HAVAS, D., BECKER, R. & RINCK, M. (2005). Grounding Language in Bodily States: The Case for Emotion. In: *The Grounding of Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (R. Zwaan & D. Pecher eds.) for discussion. See also BHALLA, M. & PROFFITT, D. R. (1999). Visual–Motor Recalibration in Geographical Slant Perception. *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 1076; WITT, J. K. & SUGOVIC, M. (2010). Performance and Ease Influence Perceived Speed. *Perception*, **39**, 1341–1353; JOSTMANN, N. B., LAKENS, D. & SCHUBERT, T. W. (2009). Weight as an Embodiment of Importance. *Psychological Science*, **20**, 1169–1174.

<sup>16</sup> PINKER, S. (2003). *The Blank Slate: The Modern Denial of Human Nature*. Penguin, London. See also TALEB, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable*. Random House, New York and Penguin, London.

<sup>17</sup> See TVERSKY, A. & KAHNEMAN, D. (1974). Judgment and Uncertainty: Heuristics and Biases. *Science* **185**, 1124–1131 for the original manuscripts and KAHNEMAN, *Thinking Fast and Slow*. Farrar, Straus and Giroux, New York for a comprehensive summary of this work.

Some of the cognitive biases identified through research include the following:

- *Confirmation bias* refers to our tendency to search for and interpret information that confirms our prior beliefs<sup>18</sup> (e.g., those who believe that arthritis pain is influenced by the weather will notice their pain more during extreme weather events, but may pay less attention when the weather is fine).<sup>19</sup>
- When we know the outcome of an event, but try to act as if we do not, we tend to be influenced by knowledge of the outcome. This tendency is known as *hindsight bias* or the *knew-it-all-along effect*.<sup>20</sup>
- The *anchoring effect* describes our tendency to rely on the first piece of information offered (the ‘anchor’) when making decisions. Subsequent judgments are influenced by initial information or beliefs (including when that information is unreliable or even arbitrary—e.g., a number on a roulette wheel or drawn from a hat).<sup>21</sup>
- A preference to remain in the same state rather than taking a risk by moving to another state is called the *status-quo bias*<sup>22</sup> (e.g., committing to the usual brand at the supermarket rather than risking an alternative<sup>23</sup>).
- Our perception of order in random sequences of coin tosses<sup>24</sup> and stock market prices<sup>25</sup> are examples of the *gambler’s fallacy*.<sup>26</sup>
- We tend to underestimate how much time we need to complete a task, even when our experience with very similar tasks suggests otherwise<sup>27</sup> (e.g., in 1957, the initial plans for the Sydney Opera House proposed opening in 1963 at a cost of \$7 million; a scaled-down version was opened in 1973 at a cost of \$102 million<sup>28</sup>). This phenomenon is aptly named the *planning fallacy*.<sup>29</sup>

<sup>18</sup> NICKERSON, R.S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, **2**, 175.

<sup>19</sup> REDELMEIER, D. A. & TVERSKY, A. (1996). On the Belief that Arthritis Pain is Related to the Weather. *Proceedings of the National Academy of Sciences*, **93**, 2895–2896. See also: ABELL, G. O. & GREENSPAN, B. (1979). Human Births and the Phase of the Moon. *The New England Journal of Medicine*, **300**, 96, for an example where people take more notice of hospital admissions during a full moon compared to other nights of the month. More generally, see BARKER BAUSELL, R. (2007). *Snake Oil Science: The Truth about Complementary and Alternative Medicine*. Oxford University Press, New York.

<sup>20</sup> FISCHHOFF, B. & BEYTH, R. (1975). I Knew It would Happen: Remembered Probabilities of Once—Future Things. *Organizational Behavior and Human Performance*, **13**, 1–16; ROESE, N. J. & VOHS, K. D. (2012). Hindsight Bias. *Perspectives on Psychological Science*, **7**, 411–426, for a review.

<sup>21</sup> TVERSKY, A. & KAHNEMAN, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, **185**, 1124–1131; CERVONE, D. & PEAKE, P. K. (1986). Anchoring, Efficacy, and Action: The Influence of Judgmental Heuristics on Self-Efficacy Judgments and Behavior. *Journal of Personality and Social Psychology*, **50**, 492.

<sup>22</sup> SAMUELSON, W. & ZECKHAUSER, R. (1988). Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty*, **1**, 7–59.

<sup>23</sup> DHAR, R. (1997). Consumer Preference for a No-Choice Option. *Journal of Consumer Research*, **24**, 215–231.

<sup>24</sup> BAR-HILLEL, M. & WAGENAAR, W. A. (1991). The Perception of Randomness. *Advances in Applied Mathematics*, **12**, 428–454.

<sup>25</sup> HUBER, J., KIRCHLER, M. & STÖCKL, T. (2010). The Hot Hand Belief and the Gambler’s Fallacy in Investment Decisions under Risk. *Theory & Decision*, **68**, 445–462.

<sup>26</sup> TVERSKY, A. & KAHNEMAN, D. (1971). Belief in the Law of Small Numbers. *Psychological Bulletin*, **76**, 105.

<sup>27</sup> KAHNEMAN, D. & TVERSKY, A. (1979). Intuitive Predictions: Biases and Corrective Procedures. *TIMS Studies in Management Sciences*, **12**, 313–327. For a review, see BUEHLER, R., GRIFFIN, D. & ROSS, M. (2002). Inside the Planning Fallacy: The Causes and Consequences of Optimistic Time Predictions. In: *Heuristics and Biases: The Psychology of Intuitive Judgment* (T. D. Gilovich, D. W. Griffin & D. Kahneman eds.). Cambridge University Press, New York.

<sup>28</sup> BUEHLER, R., GRIFFIN, D. & ROSS, M. (1994). Exploring the “Planning Fallacy”: Why People Underestimate their Task Completion Times. *Journal of Personality and Social Psychology*, **67**, 366.

<sup>29</sup> See e.g. FLYVBJERG, B. (2003). *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge University Press, Cambridge, UK.

Sometimes our basic pattern recognition abilities are shaped by very specific knowledge and expectations. For example, a ‘backward message’ attributed to rock music is easy to hear when you know (through direct suggestion) what phrase to listen for. The effects of expectancy can make these ‘hidden messages’ (or features) in ambiguous auditory or visual signals seem intelligible even when there is no actual message or feature.<sup>30</sup> By knowing specifically what to look or listen for, or what to remember, some characteristics that are consistent with a specific piece of information are ‘sharpened’, exaggerated, and emphasized, whereas other characteristics that are inconsistent, are ‘levelled’, toned down, and weakened.<sup>31</sup> This general process of sharpening and levelling information that is consistent with one’s expectations can be referred to generically as an *expectancy effect*.<sup>32</sup>

Our ability to recognize patterns is also influenced by *general knowledge* and expectations. Under most circumstances, using language as an example, we are unaware of our reliance on syntax, word frequency, our background knowledge of specific individuals or the topic of conversation. Under ‘noisy’ conditions (e.g., reading untidy handwriting, listening to a telephone adjacent to a busy road, tracking an unfamiliar accent or watching a badly distorted video), the pace seems fast, but our general knowledge allows us to fill in the gaps and resolve ambiguity. As the ‘noise’ increases, our abilities to fill gaps tend to deteriorate making us more prone to mistakes.

The reliance on specific and general knowledge is what makes it difficult for machines to perceive stimuli and complete some tasks that humans find trivial.<sup>33</sup> Computer scientists and researchers in machine learning (e.g. A.I.) can attest to the incredible amount of stored information and processing required for a computer to interpret seemingly simple information such as a handwritten post code on an envelope.<sup>34</sup> Subjectively, however, we just open our eyes and apprehend it because we are all experts in coping with the normal stimuli of our everyday lives. We are not even aware of having made an interpretation, and we take for granted the many cognitive processes that lurk beneath the surface of our external behaviour when *we* are the experts.

To non-experts, the cognitive feats that specialists are capable of performing often seem impressive, even extraordinary. For example, chess masters can play the game blindfolded with high levels of performance,<sup>35</sup> and can remember the exact configuration of pieces on a chessboard after only a few seconds of exposure (about 93% correct for configurations of about 25 pieces<sup>36</sup>); radiologists can detect and name 70% of abnormalities in chest X-rays after seeing them for only 200 ms,<sup>37</sup> and experienced waiters can memorize orders for up to 16 people without taking notes, and while engaging in unrelated conversation.<sup>38</sup>

<sup>30</sup> VOKEY, J. D. & READ, J. D. (1985). Subliminal Messages: Between the Devil and the Media. *American Psychologist*, **40**, 1231–1239. The New Zealand case of *Bain* raised a particularly striking example of this problem: *Bain v. The Queen* [2009] NZSC 16.

<sup>31</sup> GIBSON, J. J. (1929). The Reproduction of Visually Perceived Forms. *Journal of Experimental Psychology*, **12**, 1–39.

<sup>32</sup> Gilovich, T. (1991). *How We know What isn't So: The Fallibility of Human Reason in Everyday Life*. Free Press, New York.

<sup>33</sup> See COLLINS, H. (1990). *Artificial Experts: Social Knowledge and Intelligent Machines*. MIT Press, Cambridge, MA. DREYFUSS, H. (1972) *What Computers Can't Do*. MIT Press, Cambridge, MA.

<sup>34</sup> LE CUN, Y., JACKEL, L. D., BOSER, B., DENKER, J. S., GRAF, H. P., GUYON, I. & HUBBARD, W. (1989). Handwritten Digit Recognition: Applications of Neural Network Chips and Automatic Learning. *Communications Magazine, IEEE*, **27**, 41–46.

<sup>35</sup> HOLDING, D. H. (1985). *The Psychology of Chess Skill*. L. Erlbaum Assoc.

<sup>36</sup> DE GROOT, A. D. (1978). *Thought and Choice in Chess*, 2nd edn. Mouton, The Hague (originally published 1946).

<sup>37</sup> KUNDEL, H. L. & NODINE, C. F. (1975). Interpreting Chest Radiographs without Visual Search. *Radiology*, **116**, 527–532.

<sup>38</sup> ERICSSON, K. A. & POLSON, P. G. (1988). Memory for Restaurant Orders. In: *The Nature of Expertise* (M. Chi, R. Glaser & M. Farr eds.). Erlbaum, Hillsdale, NJ; ERICSSON, K. A. & POLSON, P. G. (1988). An Experimental Analysis of the Mechanisms of a Memory Skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 305; THOMPSON, M. B., TANGEN, J. M.

According to the well-established *exemplar theory* of categorization,<sup>39</sup> the identification of category members in everyday classification (e.g., a bird, a table, or a car) or expert classification (e.g., an abnormal chest X-ray, a patient with myocardial ischaemia, or a poor chess move) is effortless because experts have acquired a large number of exemplars. They respond to new items by reference to their similarity to those previously encountered. Often this sensitivity develops effortlessly and without any intention to learn structures or categories.<sup>40</sup> Yet, both common experience and laboratory research alike demonstrate that this tacit sensitivity influences performance and expectations in virtually every task we undertake.<sup>41</sup> Our perception and cognition—the way we see, hear and remember the world—is, in a very real sense, shaped by the sum of our experiences.

Of course, the acquisition of expertise requires more than just an accumulation of experiences; it requires good quality and timely feedback.<sup>42</sup> A ballet dancer practicing a plie or pirouette in front of a mirror can see immediately which aspects of her posture need correcting and, with sufficient practice and instruction, is able to refine even the most difficult and unnatural of movements. There are situations within policing where feedback is similarly informative and immediate. In acquiring firearm and marksmanship skills, for example, police recruits learn through training how slight changes in position and breathing can affect accuracy and consistency between shots. Recruits see immediately where they have hit the target.

Our sensitivity to feedback and cues in our environment do not always lead to desirable outcomes. In some environments, we can use or ‘learn’ from the wrong cues. Robin Hogarth describes these as ‘wicked’ environments and provides the example of a physician in the early 20th century who reportedly developed a technique to diagnose patients with typhoid by palpating their tongues with his unwashed hands.<sup>43</sup> When each of his patients fell ill with typhoid, he mistakenly took this as positive feedback that his intuitions and method were correct, when in fact he was simply transferring typhoid from one patient to the next. Environments that are less structured and regular than ballet or the firing range may create an *illusion of skill* where we judge our abilities to be more pronounced than they actually are.<sup>44</sup> The stock market is a good example of an irregular environment, where it has not proved possible to make accurate predictions about stock prices consistently. Nevertheless, professional investors and fund managers routinely make predictions about the future of stock prices, and despite their confidence in these predictions (on average), regularly fail to outperform the market—demonstrating a level of performance more akin to chance (and for many individual investors, often worse

& SEARSTON, R. A. (2014). Understanding expertise and non-analytic cognition in fingerprint discriminations made by humans. *Frontiers in Psychology*, **5**, doi: 10.3389/fpsyg.2014.00737.

<sup>39</sup> BROOKS, L.R. (1978). Nonanalytic Concept Formation and Memory for Instances. In: *Cognition and Categorisation* (E. Rosch & B. Lloyd eds.). Erlbaum, Hillsdale, NJ.

<sup>40</sup> HOGARTH, R. (2010). *Educating Intuition*. Chicago University Press, Chicago.

<sup>41</sup> NORMAN, G., YOUNG, M. & BROOKS, L. (2007). Non-Analytical Models of Clinical Reasoning: the Role of Experience. *Medical Education*, **41**, 1140–1145.

<sup>42</sup> KAHNEMAN, *Thinking Fast and Slow*.

<sup>43</sup> HOGARTH, R.M. (2008). On the Learning of Intuition. In: *Intuition in Judgment and Decision Making* (H. Plessner, C. Betsch & T. Betsch eds.). Lawrence Erlbaum Associates, Inc, Mahwah, NJ.; THOMAS, L. (1983). *The Youngest Science: Notes of a Medicine Watcher*. Viking, New York, p. 22.

<sup>44</sup> KAHNEMAN, *Thinking Fast and Slow*. This might help explain the persistence of ‘fields’ and individuals purporting to be experts in forensic domains where there is no (evidence of) actual expertise.



than chance).<sup>45</sup> Without regularity or structure in the learning environment, no amount of practice and feedback will generate genuine expertise.

As a result of the automaticity of context effects, we tend to believe that the information we receive through our senses is an accurate reflection of the world, uncontaminated by our preferences, preconceptions, beliefs and interpretations.<sup>46</sup> This *naïve realist* view of perception and cognition is mistaken and potentially misleading.<sup>47</sup> The inability to recognize the extent to which prior experience shapes our judgments and decisions has been labelled the ‘curse of knowledge’.<sup>48</sup> When you know something, it is difficult to put yourself in the shoes of someone who does not know it. Moreover, we tend to believe that other people are influenced by prior experience, but that we are not—sometimes described as *bias blindness*, a *bias blind spot*,<sup>49</sup> or more informally, the *not me fallacy*.<sup>50</sup> Our beliefs and experiences automatically influence our judgments, with little or no effort, and with no sense of voluntary control.<sup>51</sup> These effects are not a sign of weakness and cannot be willed away; just as we cannot use willpower to overcome the impressions created by most visual illusions—recall Figure 1.

## 2.2 Contextual bias in the forensic sciences

Forensic science evidence presented in court is often neatly packaged as an independent source of evidence in the form of a detailed report and/or testimony from an impartial and experienced expert. Confronted with the end-product, it is difficult to appreciate the many steps involved in producing this evidence—from the time trace evidence is left at a crime scene, through collection, processing, analysis, interpretation (and verification) and preparation for presentation in court.

The process in which trace evidence is recovered involves many people with diverse experience and backgrounds and varying levels and types of expertise, preconceptions and knowledge about the case. Upon arrival at a crime scene, first responders and investigators are often faced with limited and unconfirmed information about the (alleged) crime. They are required to reconstruct events and fill in gaps with preliminary notifier reports, witness statements and impressions, usually under considerable time and resource pressures. Quite often, chance plays a major role in the detection of evidence. The types of specimens collected are also dependent on the tools and training available to investigators, the kinds of analytical facilities and resources available, and the type of crime. Ultimately, the chance that

<sup>45</sup> BOGLE, J.C. (2000). *Common Sense on Mutual Funds: New Imperatives for the Intelligent Investor*. Wiley, New York; GRINBLATT, M. & TITMAN, S. (1992). The Persistence of Mutual Fund Performance. *Journal of Finance*, **42**, 1977–1984; ELTON, E. J., *et al.* (1997). The Persistence of Risk-Adjusted Mutual Fund Performance. *Journal of Business*, **52**, 1–33; and KAHNEMAN, *Thinking Fast and Slow*, pp. 212–216 for a more expansive review and further examples of the illusion of skill in stock picking and in other professional fields. See also BARBER, B. M. and ODEAN, T. (2002). Trading is Hazardous to Your Wealth: the Common Stock Investment Performance of Individual Investors. *Journal of Finance*, **55**, 773–806.

<sup>46</sup> SEGALL, M., CAMPBELL, D. & HERSKOVITZ, M. (1966). *The Influence of Culture on Visual Perception*. Bobbs-Merrill, New York.

<sup>47</sup> ROSS, L. & WARD, A. (1995). Psychological Barriers to Dispute Resolution. *Advances in Experimental Social Psychology*, **27**, 255–303.

<sup>48</sup> CAMERER, C., LOEWENSTEIN, G. & WEBER, M. (1989). The Curse of Knowledge in Economic Settings: An Experimental Analysis. *The Journal of Political Economy*, **97**, 1232–1254.

<sup>49</sup> PRONIN, E., GILOVICH, T. & ROSS, L. (2004). Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self versus Others. *Psychological Review*, **111**, 781–799.

<sup>50</sup> LILJENFELD, S.O., AMMIRATI, R. & LANDFIELD, K. (2009). Giving Debiasing Away: Can Psychological Research on Correcting Cognitive Errors Promote Human Welfare? *Perspectives on Psychological Science*, **4**, 390–398.

<sup>51</sup> NISBETT, R. E. & WILSON, T. D. (1977). Telling More than We can Know: Verbal Reports on Mental Processes. *Psychological Review*, **84**, 231–259.

one piece of evidence is collected, as opposed to others, is driven by the perspectives and preconceptions of investigators, organizational expectations and capacities, and the information available.

One of the dangers in any investigation is the failure to consider (or pursue) alternative possibilities. There is a real risk that investigators and forensic analysts will engage in *tunnel vision*.<sup>52</sup> This is where, usually unwittingly, investigators interpret evidence in a manner that tends to confirm the basic case theory. Tunnel vision—which is related to sharpening and levelling, expectancy effects and confirmation bias—illustrates the human tendency to hold onto previously formed theories, hypotheses or beliefs even in the face of new or disconfirming evidence. Interpretive flexibility, in conjunction with the potential for other information (e.g. beliefs, suspicions and other evidence) to cue, may result in traces, data and readouts being interpreted in ways that are consistent with expectations (and in ways that are different to how the same data might have been interpreted if the investigator or forensic analyst was shielded from gratuitous information, had different assumptions, or was exposed to the evidence in a different sequence).<sup>53</sup> Although some failures and risks might be avoided, some cannot be. For example, the assumptions, expectations and prior experiences of the investigator may bias the interpretation of a crime scene and the collection of evidence.

Beyond the crime scene, it is a *human* analyst who is required to sort traces and samples, select and run tests or undertake analyses in order to report on the findings. Almost all forensic science and medicine techniques rely upon input and interpretation by analysts, often in order to link a trace or sample to a particular person or source.<sup>54</sup> Some types of comparison or analysis rely on tools and/or technologies to assist with the interpretation (e.g., a fingerprint database or image enhancement), but the final decision almost always rests with the analyst. Ideally, the distinction between traces that originate from the same source and those that originate from different sources would be obvious. In reality, traces are regularly degraded (e.g., smudged or partial, mixed or poorly resolved) or quite similar to the reference (e.g., the suspect's profile), resulting in ambiguity and increasing scope for erroneous interpretations.

There are a number of obvious ways in which contextual cues, observer and expectancy effects, as well as anchoring and priming might influence the process of evaluating and interpreting evidence. If someone is affected by domain-irrelevant information, they respond differently than they would if they were not exposed to this extraneous information. This information might relate to the trace, the suspect, or the case.<sup>55</sup> One study by Dror *et al.* demonstrated how exposure to domain-irrelevant information led a small sample of latent fingerprint examiners to contradict their own opinions as to whether the same two prints matched.<sup>56</sup> Dror *et al.* followed up this study, by showing that other forms of contextual information (e.g., being told that the suspect confessed or emotion-evoking case information) could also influence analyses by both novice and experienced fingerprint examiners.<sup>57</sup> They

<sup>52</sup> FINDLEY, K. A. & SCOTT, M. S. (2006). Multiple Dimensions of Tunnel Vision in Criminal Cases. *Wisconsin Law Review*, **2**, 291–397.

<sup>53</sup> KRANE, D. *et al.* (2008). Sequential Unmasking: a Means of Minimizing Observer Effects in Forensic DNA Interpretation. *Journal of Forensic Science*, **53**, 1006.

<sup>54</sup> TANGEN, J. M. (2013). Identification Personified. *Australian Journal of Forensic Sciences*, **45**, 315.

<sup>55</sup> THOMPSON, W. C. (2011). What Role should Investigative Facts play in the Evaluation of Scientific Evidence? *Australian Journal of Forensic Sciences*, **43**, 123–134.

<sup>56</sup> DROR, I. E., CHARLTON, D., and PÉRON, A. E. (2005). Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications. *Forensic Science International*, **156**, 74–78.

<sup>57</sup> DROR, I. E. & CHARLTON, D. (2006). Why Experts make Errors. *Journal of Forensic Identification*, **56**, 600–616; 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**, 799–809.

also found that the interpretation of complex DNA profiles (e.g. mixtures), as well as forensic anthropological assessments of gender, age and race, were vulnerable to similar extraneous influences.<sup>58</sup>

Our concern is with the information and conditions that adversely affect the analyst's perception and interpretation. Scientific research on confirmation bias suggests that ambiguous information that is consistent with what is expected tends to be sharpened, whereas ambiguous information that is inconsistent with what is expected tends to be levelled.<sup>59</sup> The extent to which information is regarded as ambiguous presumably varies between individuals and the same individual's performance may vary diachronically.<sup>60</sup> Hence, ambiguous information may not seem ambiguous as a result of confirmation bias and is much in the 'eye of the beholder'. The extent to which these sharpening and levelling processes will affect the examiner's interpretation also depends on a number of other factors, including the amount of actual ambiguity—which can be molded to fit one's expectations. If the information in the sample is unambiguous (e.g., the blood type test reads O negative), then it is almost impossible for sharpening or levelling to change the perception and interpretation (to make it AB positive). If, however, the information is ambiguous (e.g. a small amount of DNA, a degraded or mixed sample, allelic drop-ins and drop-outs, stutters, peak-height imbalances), then certain bits of this ambiguous information may seem more relevant than others in light of extraneous information (e.g. a suspect's profile).<sup>61</sup> Evidence found at crime scene is often of degraded quality and quantity, includes noise and distortions, and other factors which often make it ambiguous.

Many forensic analysts are exposed to potentially corrosive domain-irrelevant information in their day-to-day work.<sup>62</sup> Not infrequently, forensic analysts are told about other strands of (potential) evidence, the investigator's background knowledge of the case and the suspect (e.g. criminal history or admissions), as well as results from other forensic analyses. Given the inevitability of errors and biases in human interpretation, it seems reasonable to assume that forensic analysts, like the rest of us, are susceptible to expectancy effects and cognitive biases. Where there is vulnerability, the only way to avoid being influenced inappropriately is to restrict access to domain-irrelevant information (or, more precisely, information with the potential to mislead). Thompson and Dror propose the separation of a forensic scientist's role into either a 'case manager'—who communicates with investigators, may help to decide what specimens to collect from the crime scene, and manages the workflow and tasks assigned in the laboratory (e.g., what to sample and which assays to run)—or an 'analyst'—who performs analyses (e.g., comparisons of trace evidence) according to instructions.<sup>63</sup> This separation of

<sup>58</sup> DROR, I. E. & HAMPIKIAN, G. (2011). Subjectivity and Bias in Forensic DNA Mixture Interpretation. *Science & Justice*, **51**, 204–208; THOMPSON, W. (2009). Painting the Target around the Matching Profile: The Texas Sharpshooter Fallacy in Forensic DNA Interpretation. *Law, Probability & Risk*, **8**, 257; NAKHAEIZADEH, S., DROR, I. E., MORGAN, R. (2014). Cognitive Bias in Forensic Anthropology: Visual Assessments of Skeletal Remains is Susceptible to Confirmation Bias. *Science & Justice*, **54**, 208–214.

<sup>59</sup> GILOVICH, T. (1991). *How We know What isn't so: The Fallibility of Human Reason in Everyday Life*. Free Press, New York.

<sup>60</sup> See 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**, 10–17 for evidence of performance changing from one time to the next.

<sup>61</sup> THOMPSON, W. C. (2013). Forensic DNA Evidence: The Myth of Infallibility. In: *Genetic Explanations: Sense and Nonsense* (S. Krimsky & J. Gruber eds.). Harvard University Press, Cambridge, MA.

<sup>62</sup> See BUTT, L. (2013). The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions: Commentary by a Forensic Examiner. *Journal of Applied Research in Memory & Cognition*, **2**, 59–60, for an analyst's perspective on the case information available to the analyst.

<sup>63</sup> THOMPSON, W. C. (2011). What Role should Investigative Facts Play in the Evaluation of Scientific Evidence? *Australian Journal of Forensic Sciences*, **43**, 123–134; DROR, I. E. (2014). Practical Solutions to Cognitive and Human Factor Challenges in Forensic Science. *Forensic Science Policy & Management*, **4**, 105–113.

roles facilitates blind analysis while allowing analysts to have access to appropriate information thereby ensuring the case manager (and the institution) is informed about the overall case.<sup>64</sup>

Another proposal involves ‘sequential unmasking’, where information is gradually revealed to the analyst.<sup>65</sup> For example, an analyst might conduct an initial examination of the trace evidence and limit their interpretation to the legible or salient parts of the sample before comparing it to the suspect (e.g., recording the possible genotypes of all possible contributors to a mixed DNA sample before learning about the profile of the victim and any suspects). Of course, the process will depend on the type of trace evidence, but the basic idea is to keep the analyst blind to potentially biasing information for as long as possible, and for the analyst to document retrospective adjustments to the interpretation in the report.<sup>66</sup> The effectiveness of these approaches and the circumstances in which blinding is considered to be necessary can really only be refined through further research pinpointing the specific types of case information that may be harmful to performance (e.g., information that reduces the overall accuracy of analysts decisions) and the kinds of traces and samples that might, in contrast to the blood typing example, represent genuine threats to interpretation. While threats to cognition are ubiquitous, not every technique and not all samples will require blinding. Nevertheless, as Section 2.3 affirms, at this point in time, there is insufficient evidence for us to ignore or discount the threat to most types of interpretation.

A range of additional, sometimes subtle ideological, factors may influence, consciously or otherwise, the conduct and performance of forensic analysts. For example, a recent study showed that forensic examiners are biased by knowledge of the side who retained their service in such a way that identical evidence is interpreted differently depending on whether they think they are working for the defense or for the prosecution.<sup>67</sup> Another example is that analysts might be outraged by the crime such that they want to put the ‘guilty’ suspect behind bars. Analysts may be more or less interested in crimes against persons of particular ethnic groups, gender, social classes or employment groups (e.g. prostitutes). Analysts might draw upon institutional philosophies (such as the need to ‘think dirty’) or pervasive beliefs—such that multiple infant deaths in the one family are compelling evidence of child abuse and even homicide (e.g. Meadow’s law).<sup>68</sup> There may be public, media and political, as well as institutional pressure to identify an offender. In serious cases, analysts might be subjected to work environments (e.g. long hours) that adversely influence performance or make it difficult to devote the necessary time and resources to other types of investigations, particularly high volume crimes.<sup>69</sup>

Many forensic science laboratories are faced with organizational constraints that, by design, create conditions that are likely to increase the chances of analysts encountering domain-irrelevant

<sup>64</sup> Some information, although relevant, may have biasing effects. Such cases are more complex. If information is not relevant, then examiners need not have it, which is relatively straightforward. A case in which the information is relevant but also biasing, however, is tricky. In such cases, the relative importance of the information ought to be weighted against the potential biasing effects. See DROR, I. E. (2012). Combating Bias: The Next Step in Fighting Cognitive and Psychological Contamination. *Journal of Forensic Sciences*, **57**, 276–277.

<sup>65</sup> See KRANE, D., et al. (2008). Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation. *Journal of Forensic Science*, **53**, 1006.

<sup>66</sup> ROSENTHAL, R. & RUBIN, D.B. (1978). Interpersonal Expectancy Effects: The First 345 Studies. *The Behavioural and Brain Sciences*, **3**, 377–386.

<sup>67</sup> MURRIE, D. C., BOCCACCINI, M.T., GUARNERA, L. A., & RUFINO, K. A. (2013). Are Forensic Experts Biased by the Side that Retained Them? *Psychological Science*, **24**, 1889–1897.

<sup>68</sup> GOUDGE, S. (2008). *Inquiry into Pediatric Forensic Pathology*. Queens Printer, Toronto; CUNLIFFE, E. (2011). *Murder, Medicine and Motherhood*. Hart Publishing, Oxford.

<sup>69</sup> See Expert Working Group on Human Factors in Latent Print Analysis. (2012). *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach*. National Institute of Standards and Technology, Washington, DC (hereafter EWG, *Latent Print Examination and Human Factors*). Dror was a member of this working group.

information. For example, in some jurisdictions, forensic analysts across several disciplines (e.g., image analysts and fingerprint examiners) share the same management hierarchy, belong to the same team or workgroup, and in some instances share office facilities and workspace. Such environments can make it difficult for analysts to avoid even inadvertent exposure to domain-irrelevant information. Some analysts are involved in the end-to-end collection and processing of evidence (e.g. many fire investigators)—thereby exposing them to the crime scene, other investigators and additional aspects of the case. Such institutional arrangements threaten the independence and, potentially, the value of resulting interpretations.

There have been endogenous attempts to reduce errors. Some forensic analysts have modified their procedures in an attempt to catch problems at the ‘back end’, before results are formally reported. One such attempt is ‘verification’. Such internal safeguards are not always well-suited to enhancing performance (e.g. by reducing the impact of context) or identifying subtle influences and errors. With verification, for example, the second analyst may not be blinded to (i.e., they are aware of) the outcome of the first analyst’s assessment, because they are (in)formally told or because a particular finding is suggested—e.g. where ‘no match’ cases (i.e. the first analyst has not declared an identification) are subject to verification. Knowing the outcome of an earlier assessment, or the beliefs of other investigators, is likely to influence an analyst (recall confirmation bias and anchoring effects) regardless of how hard she tries to resist.<sup>70</sup> Yet, such verification (or peer review, as it is sometimes styled) is commonly presented, and accepted, as a guarantee of the reliability of interpretations and opinions.<sup>71</sup>

The organization of many forensic science institutions and their workflows unnecessarily exposes the humans working in them, particularly those involved in interpretation and analysis, to real threats from contextual influences and contamination.<sup>72</sup> While the impact of these effects on the vast majority of techniques is unknown, preliminary studies (e.g. research by Dror) suggest that threats from contextual influences, with the potential to erode the probative value of expert evidence, including evidence derived using techniques that are (otherwise) demonstrably reliable, are real and warrant institutional responses.<sup>73</sup>

### 2.3 *Authoritative reviews of the forensic sciences*

Several recent reviews reinforce the orthodox nature of our concerns. In the remainder of this introduction, we succinctly advert to recent recommendations by peak scientific bodies and independent judicial inquiries.

In 2006, following congressional appropriation, the U.S. National Academy of Sciences (NAS) formed a multidisciplinary National Research Council (NRC) committee to review the condition of the

<sup>70</sup> In the Mayfield misattribution, two senior analysts ‘verified’ the mis-identification.

<sup>71</sup> On peer review and its limitations, see EDMOND, G. (2008). Judging the Scientific and Medical Literature: Some Legal Implications of Changes to Biomedical Research and Publication. *Oxford Journal of Legal Studies*, **28**, 523.

<sup>72</sup> For a review of the scientific literature regarding contextual influences in forensic science, see KASSIN, S., DROR, I.E. & KUKUCKA, J. (2013). The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions. *Journal of Applied Research in Memory and Cognition*, **2**, 4–52.

<sup>73</sup> The existence of bias does not mean that the conclusion is erroneous. There may be a need to distinguish between the decision process and the decision outcome (forensic conclusion). Bias affects the decision-making process. Whether it alters the final conclusion depends on the level and direction of the bias, as well as the difficulty of the forensic comparisons (as they become more difficult—closer to the decision threshold, the more likely bias will affect the final conclusion). However, even in cases in which bias does not alter the decision outcome, it still affects examiners’ confidence in the conclusion and how it is presented in court. See DROR, I. E. (2009). On Proper Research and Understanding of the Interplay between Bias and Decision Outcomes. *Forensic Science International*, **191**, 17–18.

forensic sciences. The result of that inquiry, ‘Strengthening the forensic sciences in the United States: A path forward (2009)’, was remarkably critical in tone.<sup>74</sup> To its surprise, the Committee encountered serious problems across the forensic sciences and expressed doubts about the evidentiary value of some forensic science techniques in regular use, particularly the non-DNA identification sciences.<sup>75</sup> Specifically addressing the issue of bias and the need for research, the NRC committee insisted that:

a body of research is required to establish the limits and measures of performance and to address the impact of sources of variability and potential bias. Such research is sorely needed, but it seems to be lacking in most of the forensic disciplines that rely on subjective assessments of matching characteristics. These disciplines need to develop rigorous protocols to guide these subjective interpretations and pursue equally rigorous research and evaluation programs. The development of such research programs can benefit significantly from other areas, notably from the large body of research on the evaluation of observer performance in diagnostic medicine and from the findings of cognitive psychology on the potential for bias and error in human observers.<sup>76</sup>

The NAS report further recommended establishing a National Institute of Forensic Science (NIFS) that, in addition to sponsoring and supervising validation studies, determining error rates, developing empirically driven standards and probabilistic forms of reporting results, would address contextual bias and threats to interpretations of evidence through attention to psychological research and revised procedures. The Committee recommended research on bias and the reform of institutional procedures and workflows.<sup>77</sup> Recommendation 5, for example, states:

The National Institute of Forensic Science (NIFS) should encourage research programs on human observer bias and sources of human error in forensic examinations. Such programs might include studies to determine the effects of contextual bias in forensic practice (e.g., studies to determine whether and to what extent the results of forensic analyses are influenced by knowledge regarding the background of the suspect and the investigator’s theory of the case). In addition, research on sources of human error should be closely linked with research conducted to quantify and characterize the amount of error. Based on the results of these studies, and in consultation with its advisory board, NIFS should develop standard operating procedures (that will lay the foundation for model protocols) to minimize, to the greatest extent reasonably possible, potential bias and sources of human error in forensic practice. These standard operating procedures should apply to all forensic analyses that may be used in litigation.<sup>78</sup>

<sup>74</sup> National Research Council. (2009). *Strengthening Forensic Science in the United States: A Path Forward*. The National Academies Press, Washington, DC. See also MNOOKIN, J., COLE, S., DROR, I. E., FISHER, B., HOUCK, M., INMAN, K. *et al.* (2011). The Need for a Research Culture in the Forensic Sciences. *UCLA Law Review*, **58**, 725–779.

<sup>75</sup> EDWARDS, H. (2010). Solving the Problems that Plague the Forensic Science Community. *Jurimetrics Journal*, **50**, 5–19. The NAS concluded that: ‘The law’s greatest dilemma in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is science in any given forensic science discipline.’ NRC, *Strengthening Forensic Science*, **87**.

<sup>76</sup> NRC, *Strengthening Forensic Science*, 8.

<sup>77</sup> For a review of several recent reports, see EDMOND, G. (forthcoming). What Lawyers should know about the Forensic “Sciences”. *Adelaide Law Review*.

<sup>78</sup> NRC, *Strengthening Forensic Science*, 24. In 2013, responsibility for reform was conferred on a committee of the National Institute of Standards and Technology (U.S.). There is no NIFS.

Risks from contextual and cognitive biases were also considered in a report jointly sponsored by the U.S. National Institute of Standards and Technology (NIST) and the National Institute of Justice (NIJ).<sup>79</sup> Assembled in the wake of the NAS inquiry, exposés around FBI mistakes and Dror's unsettling research findings, the report of an Expert Working Group (EWG) reiterated the need to attend to context and bias, though this time explicitly in response to one of the oldest and most venerated forensic techniques, namely latent fingerprint comparison.<sup>80</sup>

Notwithstanding routine admission and reliance upon the opinions of latent fingerprint examiners as positive identification evidence (i.e. individualizations), the EWG made several disruptive recommendations, including:

**Recommendation 3.3:** Procedures should be implemented to protect examiners from exposure to extraneous (domain-irrelevant) information in a case.

Simultaneously, in Scotland, Sir Anthony Campbell undertook an inquiry into the controversy emerging out of the misattribution of a latent fingerprint to P.C. Shirley McKie. Lord Campbell conducted a comprehensive review of current practice and proposed many reforms. The 'Scottish Fingerprint Inquiry' report placed conspicuous emphasis on the need to attend to contextual bias. The following were among the recommendations:

**Recommendation 6:** The SPSA [Scottish Police Services Authority] should review its procedures to reduce the risk of contextual bias.<sup>81</sup>

**Recommendation 7:** The SPSA should ensure that examiners are trained to be conscious of the risk of contextual bias.<sup>82</sup>

**Recommendation 8:** The SPSA should consider what limited information is required from the police or other sources for fingerprint examiners to carry out their work, only such information should be provided to examiners, and the information provided should be recorded.<sup>83</sup>

These later examples relate to fingerprint evidence, but many of the findings and implications have general application across the forensic sciences and beyond USA and Scotland.

For this article, the most important findings and implications from these independent and authoritative reviews are: (i) widespread recognition, particularly among experienced research scientists (and attentive others), of the potential for contextual influences and cognitive biases to create problems that require the imposition of measures to understand and (in some cases) eliminate them; (ii) the fact that even demonstrably reliable techniques, practiced by examiners with decades of experience, are susceptible to corrosive influences that threaten the probative value of the resulting opinion evidence; (iii) techniques that are less reliable than latent fingerprint comparison, or involve difficult interpretive

<sup>79</sup> EWG, *Latent Print Examination and Human Factors*. Recently, NIJ and NIST have formed an expert group of psychologists to examine human factor issues in forensic science, and to guide all the domain expert groups on how to minimize bias and other cognitive issues. See <http://www.nist.gov/forensics/osac/hfc.cfm> [Accessed 1 October 2014].

<sup>80</sup> DROR, CHARLTON and PÉRON, 'Contextual information'. See generally, COLE, S. (2001). *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Harvard University Press, Cambridge, MA; COLE, S. (2009). Forensics without Uniqueness, Conclusions without Individualization: The New Epistemology of Forensic Identification. *Law, Probability, & Risk*, **8**, 233; KOEHLER, J. & SAKS, M. (2010). Individualization Claims in Forensic Science: Still Unwarranted. *Brooklyn Law Review*, **75**, 1187.

<sup>81</sup> Lord Campbell. (2011). *The Fingerprint Inquiry Report*. APS Group Scotland, Edinburgh, Scotland, para. 35.137.

<sup>82</sup> Campbell, *The Fingerprint Inquiry Report*, para. 35.138.

<sup>83</sup> Campbell, *The Fingerprint Inquiry Report*, para. 35.139.

processes, are likely to be more susceptible to bias and suggestion induced, however (in)advertently, by information, procedures and institutional processes; (iv) experience and expertise (like awareness of dangers from contextual bias) do not necessarily insulate an analyst or enable them to overcome contextual threats; (v) in the vast majority of cases, the forensic analyst will not know if their interpretation is in fact contaminated; and (vi) many forensic science institutions have not unilaterally responded to threats posed by contextual influences, cognitive biases and cross-contamination.

The recent reports and inquiries are also revealing in their restricted purview. Focussed on the need for validation and reliability studies, research-based standards and reducing the impact of bias in the collection, analysis and reporting of incriminating opinions, they have not considered the implications of contamination and bias on the presentation and evaluation of incriminating expert evidence ‘downstream’—in interrogations, plea negotiations, trials and appeals. Moreover, the reports do not consider the promiscuity of contamination and bias. That is, the promiscuous manner in which information (and incriminating opinions) might contaminate other witnesses and evidence.

### 3. Contextual bias and cross-contamination at trial

Legal institutions have been largely inattentive to—and, therefore, ineffective in exposing or meaningfully addressing—the potential problems created by contextual influences and cognitive biases and their corrosive implications for the administration of justice. In addition, there have been very few attempts to ‘prevent or melt the snowball’: by preventing the undesirable contamination of various strands of evidence or endeavouring to expose and disentangle contamination, in order to enhance the reliability of expert evidence and facilitate its rational assessment.

#### 3.1 Admissibility gatekeeping: regulating biased incriminating expert evidence

Admissibility decisions involve determining whether evidence can be admitted and (potentially) relied upon during legal proceedings. Generally, admissibility is governed by logical relevance—the ability of the evidence to rationally influence the assessment of contested facts (e.g. FRE 401, 402 and UEL Sections 55, 56).<sup>84</sup> All adversarial jurisdictions have rules allowing those with relevant expertise (or ‘specialized knowledge’) to express opinions based on their abilities or knowledge (e.g. FRE 702 and UEL Section 79).<sup>85</sup>

In practice, most jurisdictions liberally admit incriminating expert opinion evidence.<sup>86</sup> This lax approach applies even when the evidence is derived from techniques that have not been formally evaluated and where the analyst is not shielded from domain-irrelevant information. Laxity is pervasive in jurisdictions that (purport to) direct attention to reliability following *Daubert v Merrell Dow Pharmaceuticals, Inc.* (e.g. U.S. federal courts and Canada), as well as those directing attention to the existence of a ‘field’, the qualifications or experience of the witness, and the utility of the evidence

<sup>84</sup> TWING, W. (1985). *Theories of Evidence: Bentham and Wigmore*. Stanford University Press, Stanford. On evidence law, more generally, see DENNIS, I. (2013). *The Law of Evidence*, 5th edn. Sweet & Maxwell, London; ROBERTS, P. & ZUCKERMAN, A. (2010). *Criminal Evidence*, 2nd edn. Oxford University Press, Oxford; SOPINKA, J., BRYANT, A. & LEDERMAN, S. (2009). *Law of Evidence in Canada*, 3rd edn. LexisNexis, Markham; LIGERTWOOD, A. & EDMOND, G. (2010). *Australian Evidence*, 5th edn. LexisNexis, Sydney.

<sup>85</sup> The references are to the Uniform Evidence Law (Australia) and the Federal Rules of Evidence 1975 (U.S.).

<sup>86</sup> See e.g. Law Commission of England and Wales. (2009). *Expert Evidence in Criminal Proceedings in England and Wales*. HMSO, London.



(e.g. England and Wales, Australia and New Zealand).<sup>87</sup> Some jurisdictions require expert evidence to go beyond what the trier of fact already knows—or is presumed to know. Such ‘common knowledge’ rules create problems when cognitive scientists are called, usually by criminal defendants, to proffer insights from experimental studies. Common law courts beyond USA tend to assume that juries and judges do not require this kind of assistance because of their lived experience. Psychological research is often (mis)characterized as ‘common sense’. Such responses enable judges to treat critical insights as ‘within the range of human experience’, or suggest that they can be adequately encapsulated in a judicial comment (or warning).<sup>88</sup> When proffered, in USA or elsewhere, it is not uncommon for psychological knowledge, including counterintuitive findings and insights about contextual bias and cross-contamination, to be trivialized by prosecutors or excluded by the trial judge.<sup>89</sup>

In addition to formal rules regulating the admission of expert opinion evidence, most jurisdictions grant the trial judge a discretion (sometimes imposed as an obligation) to exclude otherwise admissible evidence if it might unfairly prejudice the accused, or might waste time and resources, in relation to its probative value (e.g. FRE 403, UEL ss135, 137 and PACE s78).<sup>90</sup> These discretions, potentially applicable to all forms of evidence, require the trial judge to balance the probative value of the evidence—the extent to which it can rationally influence the assessment of contested facts—against any unfair prejudice it may create. Unfair prejudice is usually understood as the danger that the evidence might mislead, or be misunderstood or over-valued by, the trier of fact, although it might also include the inability to explore the evidence or adequately convey limitations through the course of an adversarial proceeding.

These ‘discretions’ are usually operationalized through objection. Unlike admissibility standards, the party challenging admission is required to persuade the trial judge that the probative value of the evidence is outweighed by the danger of unfair prejudice. In practice, prosecutors and judges have not been particularly attentive to probative value or some very real dangers, when considering the exercise of discretions around the admission and use of incriminating expert opinion evidence. Instead, they tend to assume that the evidence is probative because a jury might attribute value when ‘weighing’ the evidence. There is often an expressed reluctance to trespass on the prerogatives of the trier of fact. While this response might be comprehensible in relation to many forms of evidence, where the trier of fact is reasonably well positioned to make an assessment (e.g. with most ordinary witness testimony), incriminating expert evidence is arguably different.<sup>91</sup> Unlike most forms of testimony, the techniques underlying incriminating expert evidence are generally susceptible to some kind of empirical evaluation (or independent support).<sup>92</sup> That is, we can (and should) obtain an indication of ability and accuracy. Moreover, the techniques can be structured to prevent corrosive influences and reduce unnecessary errors.

The danger of unfair prejudice manifests in the danger that the jury—oblivious to expectancy effects, priming, confirmation bias, contextual influences and their implications for forensic

<sup>87</sup> *Daubert v. Merrell Dow Pharmaceuticals, Inc* 509 US 579 (1993); *R v. Trochym* [2007] 1 SCR 239. See also REDMAYNE, M. (2001). *Expert Evidence and Criminal Justice*. Oxford University Press, Oxford and EDMOND, G. (2014). *The Admissibility of Forensic Science and Medicine Evidence under the Uniform Evidence Law*. *Criminal Law Journal*, 38, 136.

<sup>88</sup> *Smith v. The Queen* (1990) 64 ALJR 588; *R v. Turner* [1975] QB 834.

<sup>89</sup> Though rarely considered, there is also a tendency to treat conscious and unconscious bias as basically the same kind of thing (see *Li v. Regina* [2003] NSWCCA 290, [71]) and to treat process issues as matters of weight rather than admissibility, see *ASIC v. Rich* [2005] NSWCA 152 [94].

<sup>90</sup> *Police and Criminal Evidence Act* 1984 (England and Wales).

<sup>91</sup> Though even this can be deceptive because eyewitness testimony is far more complicated than most assume, and even experienced investigators and judges are not particularly good at picking lies.

<sup>92</sup> NRC, *Strengthening the Forensic Sciences*, 9–10, 113, 122, 184, 189.

techniques based on human interpretation—will over-value incriminating forensic science evidence; treating incriminating expert opinions as basically reliable and independent corroboration of other inculpatory evidence. Unfair prejudice also manifests in the difficulty of explaining the risks posed by biases and cross-contamination (along with the significance of validation and error rates) to the trier of fact in a manner that will convey the magnitude of potential problems or enable them to be rationally assessed. The need to convey limitations (to jurors and judges) and their threat to proof currently falls entirely upon (relatively) poorly resourced defence lawyers. The fact that historically, and even after the NAS and other reports, so few lawyers or judges have referred to the problems created by insufficiently reliable evidence, or addressed the dangers raised by opinions that have not been shielded from contextual bias and cross-contamination, suggests that courts are by and large oblivious, or unreceptive, to the seriousness of the threat.

Admissibility jurisprudence and decision-making, including the decisions made by prosecutors, should attend to the serious threats posed by contextual bias and cross-contamination. For, exposures can be eliminated or practically managed in most circumstances. Courts should be cautious about admitting contaminated expert opinions until we are confident that contamination does not pose a genuine threat to reliability and proof.<sup>93</sup>

### 3.2 *Representing forensic science and expert evidence at trial*

Once incriminating expert evidence is admitted, in theory adversarial safeguards (most conspicuously cross-examination, jury directions and the standard of proof) protect the accused from unfairness and wrongful conviction.<sup>94</sup> In practice, it becomes the responsibility of the defence lawyer to identify, explain and successfully convey evidentiary limitations with expert evidence to the trier of fact (and an envisioned court of appeal, should the accused be convicted).

Conventional accusatorial practice and adversarial trial mechanisms (and safeguards)—such as prosecutorial restraint, cross-examination, rebuttal experts, placing restrictions on the manner in which opinions are expressed, judicial directions and the standard of proof—have consistently fallen short in affording substantial protection to the accused against speculative incriminating expert opinions and opinions obtained in ways that are insensitive (or insufficiently sensitive) to contextual bias and cross-contamination.<sup>95</sup> Because of the way trials are resourced and operationalized, ‘in practice’, trial safeguards and protections have not made trials featuring incriminating expert evidence fair, nor have they placed the trier of fact in a position conducive to the rational assessment of the expert evidence.

Perhaps, the most widely celebrated means of drawing attention to weaknesses in incriminating expert evidence is through cross-examination. In reality, the cross-examination of forensic scientists is not infrequently superficial, dependent as it almost always is on the resourcing, technical sophistication

<sup>93</sup> See KRANE, D., *et al.* (2008). Sequential Unmasking: a Means of Minimizing Observer Effects in Forensic DNA Interpretation. *Journal of Forensic Science*, **53**, 1006.

<sup>94</sup> GARRETT, B. (2011). *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*. Harvard University Press, Cambridge, MA.

<sup>95</sup> COLE, S. (2011). Splitting Hairs? Evaluating “Split Testimony” as an Approach to the Problem of Forensic Expert Evidence. *Sydney Law Review*, **33**, 459; McQUISTON-SURRETT, D. & SAKS, M. (2009). The Testimony of Forensic Identification Science: What Expert Witnesses Say and what Factfinders Hear. *Law & Human Behavior*, **33**, 436; EDMOND, G. & SAN ROQUE, M. (2012). The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial. *Current Issues in Criminal Justice*, **24**, 51.

and abilities of publicly funded defence lawyers.<sup>96</sup> Many defence lawyers focus on chain of custody and expert witness credibility rather than substantially engage with technical and methodological problems.<sup>97</sup> When it comes to cross-examination, the technical competence of lawyers (and forensic analysts) along with beliefs about jury abilities often define the effective limits to inquiry and may render cross-examination of limited value in substantially addressing issues and dangers for the trier of fact. In relation to domain-irrelevant information, the fact that many threats to cognition and interpretation are unconscious and unrecorded (and therefore not known), substantially weakens the value of cross-examination. The most that might be obtained from a forensic analyst is a concession about good practice, even though such practices are not required in many state-run forensic science laboratories and the analyst might invoke experience and personal attributes to dismiss defence questions about extraneous contextual information and influences as hypothetical, far-fetched and professionally offensive.

Where the defence is able to secure the services of their own rebuttal expert, and that witness' testimony is admitted, it seems unlikely that rebuttal evidence will moderate the reception of opinions obtained in conditions insensitive to contextual bias and cross-contamination. In contrast to shielding analysts from domain-irrelevant information, there is no evidence that methodological 'lessons' at trial are effective.<sup>98</sup> Rebuttal evidence does not function symmetrically in the criminal trial. Unlike the state-employed, experienced forensic analyst who undertook the original analysis and expresses his apparently (and implicitly) disinterested incriminating opinion at trial, rebuttal witnesses are rarely practicing forensic scientists and in many cases will not have tested or analysed the evidence from the specific case. Rather, they tend to be academics or consultants asked to comment on (usually criticize) a method or set of practices used by a state-employed forensic analyst.<sup>99</sup> Unlike state-employed analysts, they are paid for their services in the particular case. Defence rebuttal experts are relatively easy to characterize as interested (or partisan) and their methodological concerns, including concerns around the need for validation studies and to guard against notorious risks from contextual bias, and other sources of contamination, are often (mis)represented and presumably (mis)understood as theoretical or even desperate. The fact that their concerns frequently embody orthodox commitments (endorsed in the NAS and other reports) tends to be overlooked or downplayed.<sup>100</sup> At best, the trier of fact is left to apply methodological criticisms to the opinion of an apparently impartial, experienced forensic analyst in relation to the assessment of the overall case assembled against the accused.

When it comes to understanding incriminating expert evidence, judicial directions, instructions and warnings to the jury also tend to be of limited utility; dependent as they are upon what the prosecutor and defence lawyer bring before the court (in conjunction with the experience of judges).<sup>101</sup> They are, therefore, unlikely to move beyond the pro-prosecution orientation to the incriminating opinion

<sup>96</sup> It can be very effective when highly skilled and well-resourced defence lawyers, typically those engaged by wealthy (civil) defendants, are engaged. This is not how most criminal charges are processed. See e.g. the treatment of expert evidence *In the matter of Pan Pharmaceuticals Limited (in liquidation)* [2008] FCA 416.

<sup>97</sup> LYNCH, M., COLE, S., McNALLY, R. & JORDAN, K. (2008). *Truth Machine: The Contentious History of DNA Evidence*. University of Chicago Press, Chicago.

<sup>98</sup> Moreover, such 'lessons' would have to be repeated in trial after trial.

<sup>99</sup> COLE, S. (2009). A Cautionary Tale about Cautionary Tales about Intervention. *Organization*, **16**, 121–141.

<sup>100</sup> COLE, S. & EDMOND, G. (2015, forthcoming). Science without Precedent: The Impact of the NAS Report on the Admissibility and Use of Forensic Science Evidence.

<sup>101</sup> The issue is not that directions, warnings and instructions are always completely ineffective. Rather, instructions tend to be given by individuals who are not conversant with the issues. Generally, they do not put the trier of fact in a position to rationally evaluate incriminating expert evidence.

evidence, though perhaps informed by any insights generated through cross-examination and rebuttal witnesses. Significantly, prosecutors and defence lawyers (and, therefore, trial judges) almost never identify risks or address the analyst's exposure to (and the possible influence of) domain-irrelevant information, including exposure to other strands of evidence. Trial judges hardly ever explain the significance of the failure to test, incorporate limitations or guard against threats to interpretation that are notorious in other domains (e.g. biomedical research).<sup>102</sup> This means that juries, judges sitting in summary jurisdiction and appellate courts rarely hear about these dangers, and never accord them the significance they may deserve. These failures improperly transfer responsibility for demonstrating unreliability and explaining pernicious epistemic implications to the defence. They simultaneously increase the risk of wrongful conviction by discounting proof beyond reasonable doubt.

The obligations of the prosecutor (as a 'minister of justice'), cross-examination, rebuttal witnesses, judicial instructions and warnings and the standard of proof are conventionally considered to be appropriate means of regulating incriminating expert evidence, including 'shaky' evidence, at trial.<sup>103</sup> Though, as this brief sketch suggests, these safeguards and protections have not proven adequate to the task. They have not drawn problems, well known to scientists, to the attention of jurors and judges. There are no appellate decisions addressing the threats posed by contextual influences and cognitive bias to expert evidence and proof.

#### 4. Contextual bias and cross-contamination in plea bargains (and interrogations)

The dangers from unreliability, contextual bias and cross-contamination are pronounced in the use of forensic science evidence in interrogations, plea bargains and charge negotiations.

Ordinarily, the state approaches the accused before trial, through defence counsel, with a plea offer—usually with a discount to the charge and/or sentence if they accept a guilty plea. As part of the arrangement, the accused is obliged to accept, perhaps with some limited modification, a set of agreed facts as well as culpability. The negotiations leading up to this 'deal' usually involve a description of the case against the accused and often the evidence, or an outline of the evidence, assembled against him. This frequently includes some indication or summary of expert reports and test results.

Ineffective as trial safeguards often are, there are even fewer opportunities to question forensic analysts or assess the techniques and processes behind incriminating opinions relied upon when negotiating pleas (or responding to questions and assertions during interrogations). Because of time constraints, the manner in which evidence is assembled and defence (and prosecutorial) performances funded and rewarded, plea bargains tend to afford limited scope for unpacking contamination and weakness in forensic science evidence.<sup>104</sup> Very often, expert evidence supporting the state's

<sup>102</sup> For an extreme example from the physics community, see COLLINS, H. (2012). *Gravity's Ghost: Scientific Discovery in the Twenty First Century*. University of Chicago Press, Chicago.

<sup>103</sup> See e.g. *Daubert v. Merrell Dow Pharmaceuticals, Inc.* 113 S. Ct. 2786, 2798 (1993). Contrast Law Commission, *Expert evidence in Criminal Proceedings in England and Wales*, 1.20-21, 3.4 and NRC, *Strengthening the Path Forward*, 85. On difference between the US and other common law (and other) jurisdictions, see EDMOND, G., COLE, S. CUNLIFFE, E. & ROBERTS, A. (2013). Admissibility Compared: The Reception of Incriminating Expert Opinion (i.e., Forensic Science) Evidence in Four Adversarial Jurisdictions. *University of Denver Criminal Law Review*, 3, 31; EDMOND, G. & VUILLE, J. (2014). Comparing the Use of Forensic Science Evidence in Switzerland, Australia, and the United States: Transcending the Adversarial/Non-Adversarial Dichotomy. *Jurimetrics*, 54, 221–277.

<sup>104</sup> It is often in the interests of defence counsel (and even some innocent defendants) to secure plea deals, especially when confronting apparently strong cases. See generally BURNS, R. (2011). *The Death of the American Trial*. University of Chicago Press, Chicago, IL; MEDWED, D. (2012). *Prosecution Complex: America's Race to Convict and Its Impact on the Innocent*. NYU, New York.

accusation is not qualified and defects are not conceded—neither disclosed nor explained. This means that defence lawyers are in the position of having to proffer advice, and those accused of crimes required to make important decisions, in circumstances where they do not always have an accurate account of the case, particularly the value of the expert evidence, assembled against them.

### 5. Contextual bias and cross-contamination on appeal

Appellate review has been remarkably insensitive to the fact that some techniques relied upon by forensic analysts are not demonstrably reliable, that interpretive processes are not always free from gratuitous contamination by investigators (and others), and that the various strands of evidence used to satisfy the standard of proof are not always independent even though they are represented and treated as independent corroboration at trial. Largely oblivious to these problems, and their corrosive potential, appellate counsel and appellate courts tend to reproduce or reinforce mistakes and oversights from the trial.

Appellate judges are constrained in the scope of review (by rules and convention) and their general inability to unilaterally go beyond the stated grounds of appeal and materials brought before them. The conduct of the trial and the evidence adduced at trial place fairly strict parameters around the possibility and scope of appeals. The standards governing review, and the willingness or ability of courts to hear particular grounds of appeal, vary. Typically, the main grounds for appeal tend to be where a decision from the trial (usually after some kind of objection) is re-considered or reviewed by a superior court. Appellate courts often assume (or infer) that a failure to object at trial was a deliberate, indeed tactical, decision.

For our purposes, the failure to exclude or moderate incriminating expert evidence when contamination threatens the probative value is, or ought often to be, an issue for courts of appeal. Where, however, there is no objection at trial, the ability to hear the issue on appeal is usually at the discretion (or leave) of the court of appeal. Where the admission of incriminating expert evidence is raised on appeal, depending on the jurisdiction, the court of appeal may re-consider admissibility against the jurisdictional admissibility standard—*de novo* review—or may restrict itself to reviewing the propriety of the trial judge's decision. Where the review focusses on the propriety of admission, ordinarily the review is focussed on the reasonableness of the decision. In U.S. federal courts, for example, this is developed in terms of whether the trial judge abused her discretion in deeming the incriminating expert evidence admissible.<sup>105</sup> It is generally difficult for an appellant to persuade appellate judges that a trial judge's decision was manifestly unreasonable, especially when the ground relates to the admissibility and use of forensic science evidence adduced by the state. These difficulties are compounded when the technique relied upon at trial has been admitted in previous proceedings.

Even where an appellate court finds that the trial judge erred, it does not follow that the appellant is entitled to a new trial or an acquittal. Appellate courts consider the various errors from the trial individually, and in combination, to determine whether the trial has miscarried thereby rendering the conviction unsafe. Where a mistake is not considered significant, it is treated as a harmless error that does not vitiate the (safety of the) conviction. On average, errors are more likely to be considered harmless when the overall case—i.e. the evidence in support of guilt—is considered strong.<sup>106</sup> Appellate courts are not usually in a good position to assess the impact of exposure to

<sup>105</sup> See e.g. *General Electric Co. v. Joiner* 522 US 136 (1997).

<sup>106</sup> See e.g. *Murdoch v. The Queen* [2007] NTCCA 1.

domain-irrelevant information on forensic science evidence, and the difficulties are compounded where the issue was not raised at trial.

An additional problem with respect to potentially biased forensic science evidence is that it may have contaminated and corrupted other lines of evidence—the bias(ing) snowball effect. And, it may have done so surreptitiously. So, when it gets to appeal, the judges may determine that a conviction is safe, even if they purport to discount or disregard some forensic science evidence, because there is sufficient additional evidence (e.g., from eyewitnesses and other forensic analysts) to support the guilt of the appellant.

Generally, trial lawyers do not raise and courts of appeal do not consider threats posed to proof beyond reasonable doubt by cross-contamination. Appellate lawyers and judges have not endeavoured to disentangle the way that exposing a forensic scientist to information about the case might not only lead to differences in opinion (through suggestion and/or confirmation bias), but the biasing information itself might simultaneously be treated by the fact-finder as independent support for the conviction. They have not considered the manner in which contaminated forensic science evidence may (e.g. when disclosed to witnesses) support or even bolster other strands of evidence. Eyewitnesses, for example, have a demonstrated tendency to strengthen their confidence when they receive positive feedback about their (perhaps tentative) initial ‘identification’ of an offender.<sup>107</sup> Unwittingly, evidence may be counted twice (or more) and the prosecution case conceived as stronger than it actually is (or would be if the same evidence was obtained without the suggestion or contamination). Other evidence may be strengthened in response to contaminated forensic science evidence that is conceived and treated as independent. The failure to unpack the potential for influence means that contaminated evidence may be treated as independent corroboration.

Overall, appeals afford good theoretical opportunities for addressing fundamental problems with incriminating expert evidence, including the way such evidence might be (mis)represented and (mis)understood at trial. Appeals provide a process that enables judges to consider whether a forensic technique and derivative opinions are admissible, the manner in which the evidence presented was substantially (un)fair, the manner in which evidence represented and combined was (in)appropriate, and whether any of these vitiated the fairness of the process and/or accuracy of the verdict. In practice, however, the issues explored in this article are rarely, if ever, raised.

## 6. Conclusion: what should we do about contextual bias and cross-contamination?

A large proportion of our forensic science (and other) evidence is routinely, though unnecessarily, exposed to risks from contamination and error. In most cases, we do not know and cannot easily determine, whether the particular analysis and resulting evidence is corrupted. Significantly, for investigations and legal practice, many forensic scientists and most lawyers, judges and jurors are oblivious to the dangers. Our research explains why it is not an appropriate response to these dangers to suggest that they are trivial or rare or can somehow be addressed through explanation and judicial warnings at trial.

Many adversarial legal systems (notably those in England, Australia and USA) have been conspicuously concerned with ‘adversarial bias’.<sup>108</sup> That is, where an expert aligns his opinions (and perhaps himself) with the party and the party’s interests in the proceedings. This can be unconscious or

<sup>107</sup> SEMMLER, C., BREWER, N. & WELLS, G. L. (2004). Effects of Postidentification Feedback on Eyewitness Identification and Nonidentification Confidence. *Journal of Applied Psychology*, **89**, 334–346.

<sup>108</sup> See e.g. HUBER, P. (1991). *Galileo’s Revenge: Junk Science in the Courtroom*. Basic Books, New York.

unintended, but it is often portrayed as deliberate and disingenuous. Concerns about adversarial bias (also known as expert partisanship) have been most conspicuous in civil proceedings and have produced significant reform initiatives across many adversarial jurisdictions.<sup>109</sup> To the extent that adversarial bias has received attention in criminal proceedings, predominantly it has been in response to expert evidence (usually rebuttal evidence) adduced by the defence, or in the aftermath of a wrongful conviction exposé. Too little attention has been directed at the quality of the state's incriminating forensic science evidence. As the NAS and other reports recommend, there is a need to focus attention on the reliability of forensic science and medicine evidence *adduced by the state*. There is a need to respond even if it turns out, after studies have been performed, that only a small proportion of the analyses and interpretations performed by forensic sciences are vulnerable to serious risks from domain-irrelevant information and context. Analyses at risk will require procedural responses.

Defence counsel and trial safeguards have rarely identified, let alone consistently and effectively explained, the epistemic threats posed by a range of undesirable influences and processes. The widespread danger of cross-contamination—our *bias(ing) snowball effect*—is almost never raised. On those rare occasions where such deficiencies are alluded to—'raised' would imply more deliberation than is warranted—treatment tends to be perfunctory. There are few references to relevant scientific literatures and the significance of threats to evidence is not adequately explained. In consequence, lay fact-finders, whether judge or jury, are left to attribute whatever significance they deem appropriate. In most cases, the issues are not raised and so there are few reasons to believe that they play any part in the administration of the trial or fact-finding. How jurors are supposed to rationally assess evidence derived from processes that were inattentive to contextual bias and cross-contamination, especially when dangers are unlikely to be conceded by the state and its experts, or substantially raised by the defence or trial judge, is far from obvious.

It is our contention that judges should be willing to exclude expert evidence where the forensic analyst is unnecessarily exposed to gratuitous case information, or operates in an environment with obvious vulnerabilities, *unless* there is independent empirical evidence of limited risk to the accuracy of analysts' interpretations. Initially, exclusion might appear to be an extreme response. However, while the risks are real and mistakes potentially devastating, they are also open to correction. The exclusion of contaminated evidence does not mean that the opportunity to obtain relevant and reliable expert evidence is lost. Most of our concerns can be allayed by having the relevant data or analyses re-interpreted by a different analyst in conditions that insulate re-analysis from domain-irrelevant information. So, for example, another fingerprint examiner might be asked to compare two prints, or a biologist asked to identify the alleles in a mixed sample blinded to the profile of the main suspect. To the extent that the techniques relied upon are actually valid and the analyst proficient, the results of re-analyses will be more valuable than the original contaminated interpretations. These second analyses and results will be genuinely independent. Even if the results are weaker or more qualified they will, nevertheless, provide insights that reflect the actual probative value of the evidence.

Courts should be reluctant to admit incriminating opinions obtained in conditions where the analyst was unnecessarily exposed to information about the case or the suspect.

The psychological case is clear. Humans are vulnerable to a range of contextual and cognitive biases. In many cases they are not able, and this includes individuals with vast forensic experience, to think their way out of the context and its invidious and sometimes subtle forms of contamination and

<sup>109</sup> See e.g. WOOLF, L. (1996). *Access to Justice*. HMSO, London; New South Wales Law Reform Commission, *Expert Witnesses*, Report No 109 (2005); *Kumho Tire Co v. Carmichael* 526 US 137 (1999).

influence.<sup>110</sup> We should, therefore, expect our forensic analysts to protect themselves against known dangers.

Given the possibility of having the analysis re-done by the state, requiring the defence to address dangers at their own risk seems incongruous with the goal of a fair trial, an accurate outcome and the elimination of reasonable doubt. At trial, the (often impecunious) accused is not in a good position to engage in methodological debates in order to elucidate real risks that may be trivialized by prosecutors, expert witnesses, jurors and even judges. Moreover, continuing admission and legal indifference mean that these issues will need to be re-canvassed in trial after trial rather than being addressed systematically by the institutions responsible for generating the evidence. It will also (unnecessarily) depend on the honesty, memory and record keeping of forensic analysts. Courts will depend on the record of any exposure and the process documented being accurate and comprehensive, even though psychological research indicates that influences can be subtle and even unconscious. Such continuing risks make shielding the analyst during the analysis (or empirical evidence of an ability to resist) preferable to trying to reconstruct exposure and unconscious influence, months or years later during a contested proceeding. It is more efficient and effective to require forensic analysts and their institutions to impose blinding ('upstream') than to try to address or repair real risks of error at trial ('downstream', after the fact) in circumstances where we will not always know what information was conveyed, the effects of contextual and other influences, and cannot determine if an error has been made. Our approach also untethers the accused from the vagaries of individual lawyers, judges or juries, their ability to secure funding for a rebuttal expert, or the willingness of an experienced state-employed forensic analyst to concede the real risks of error and undesirability of persisting with processes that have been in place, in many cases, for decades.

The failure to address (or disclose) the risks posed by contextual bias and cross-contamination means that the value of forensic science evidence is systematically misrepresented. The failure to remove threats to analytical and interpretive processes means that the state obtains the benefit of never conceding the very real risk that an analyst made a mistake whether because of contextual bias or other human practices such as collection, mishandling, contamination or miscalculation. The state should not have such an unfair advantage, especially when mistakes are unavoidable even in well-designed systems.<sup>111</sup> The state has an obligation to use reliable expert evidence, and to develop and present that evidence as accurately as is reasonably possible.<sup>112</sup> This is particularly important in a system that relies on evidence being rationally evaluated by jurors (and judges).

Finally, this article suggests that lawyers and judges are not engaging with relevant scientific knowledge. Judges seem overly confident about the operation of their systems as well as in the forensic science evidence adduced by the state (or on behalf of the state by private corporations). The failure to

<sup>110</sup> Consider the way the impressions of experienced forensic pathologists fed into suspicions in *R v. Matthey* [2007] VSC 398 and *Gilham v. R* [2012] NSWCCA 131. See also GOUDGE, *Inquiry into Pediatric Forensic Pathology in Ontario*. Queens Printer, Toronto.

<sup>111</sup> National Academy of Sciences and the Institute of Medicine. (1999). *To Err Is Human: Building A Safer Health System*. National Academies Press, Washington, DC.

<sup>112</sup> EDMOND, G. & ROBERTS, A. (2011). Procedural Fairness, the Criminal Trial and Forensic Science and Medicine. *Sydney Law Review*, 33, 359.



engage with scientific knowledge and insights poses a threat to legal proceedings and institutional legitimacy. Ironically, exaggerated judicial confidence has parallels with the historical reluctance of some forensic scientists to recognize and publicly concede frailties with their practices.

**Funding**

Australian Research Council (ARC) (FT0992041 and LP100200142) to G.E., (LP120100063) to J.M.T. and G.E., (N41756-10-C-3382); the United States National Institute of Standards and Technology (NIST) and Federal Bureau of Investigation (FBI) to I.E.D.