Journal of Applied Research in Memory and Cognition 7 (2018) 161

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Commentary

Comment on Kukucka, Kassin, Zapf, and Dror (2017), "Cognitive Bias and Blindness: A Global Survey of Forensic Science Examiners"

William R. Oliver*

Regional Forensic Center, United States

To the Editor:

I read "Cognitive Bias and Blindness: A Global Survey of Forensic Science Examiners" by Drs. Dukucka, Kassin, Zapf, and Dror with interest (Kukucka, Kassin, Zapf, & Dror, 2017). In that article, the authors cite my research on the effect of history and context in forensic pathology diagnosis as an example of increased error when "irrelevant contextual information suggested the presence of a pattern (Oliver, 2017)." Unfortunately, this severely misinterprets my work.

In fact, the opposite is true. The referenced study suggests that access to history and context improved diagnoses by approximately 20% and significantly decreased error. Indeed, the final sentence of the article reads as follows: "This study and those of medical decision making suggest that data hiding will decrease diagnostic accuracy."

It is but one of a long line of studies over many years by many authors that demonstrate the importance of history and context in medical diagnosis, including diagnosis in pathology. Imagine my surprise to see it cited as an article that claims the opposite of what I claimed.

Conflicts of Interest Statement

The author declares no conflict of interest.

Keywords: Cognitive bias, Forensic pathology, Inference, Error

References

Kukucka, J., Kassin, S. M., Zapf, P. A., & Dror, I. E. (2017). Cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *6*(4), 452–459.

Oliver, W. R. (2017). Effects of history and context on forensic pathologist interpretation of photographs of patterned injury of the skin. *Journal of Forensic Sciences*, *62*(6), 1500–1505.

Received 4 December 2017;
accepted 15 January 2018

Author Note

* Correspondence concerning this article should be addressed to William R. Oliver, Regional Forensic Center, 2761 Sullins St, Knoxville, TN 37919, United States. Contact: oliver@billoblog.com

Journal of Applied Research in Memory and Cognition 7 (2018) 162–163

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

When Expert Decision Making Goes Wrong: Consensus, Bias, the Role of Experts, and Accuracy

Itiel E. Dror*

University College London (UCL), United Kingdom

Jeff Kukucka
Towson University, United States

Saul M. Kassin, Patricia A. Zapf
John Jay College of Criminal Justice, United States

Oliver (2017) presented a study in which contextual information impacted medical examiners' decision making. Results showed that this information increased consensus among examiners by 18%, as well as confidence in their judgments. These findings are highly consistent with myriad research showing the impact of contextual information on decision making—including among experts within established forensic science domains such as fingerprinting and DNA (for a review, see Kassin, Dror, & Kukucka, 2013).

In his commentary on "Cognitive Bias and Blindness " (Kukucka, Kassin, Zapf, & Dror, 2017), Oliver (2018) interprets his own findings as proof that contextual information increased accuracy while diminishing error. Yet increased consensus and confidence cannot be used to infer accuracy. Rather, Oliver's (2017) findings betray the fact that experts were more likely to disagree when they based their conclusions solely on the medical -relevant data. Adding nonmedical , irrelevant contextual information masks this problem by artificially increasing consensus and confidence, both of which can be misinterpreted as indicators of accuracy.

Disagreement between experts is a cause for concern. For example, DNA experts can reach different conclusions from the same data. Dror and Hampikian (2011) found that DNA experts disagreed as to whether or not a suspect could have contributed to a DNA mixture, even when these experts worked in the same lab and followed the same procedures and protocols. The National Institute of Standards and Technology (NIST) replicated this finding using DNA that was analysed with established statistical tools that are regularly used in court (Coble, 2015). Even more concerning is intra- expert disagreement, whereby the same expert reaches different conclusions from the same data examined at two different times. In one study, for example, fingerprint experts reached a different conclusion 10% of the time when they examined the same fingerprints on two separate occasions (Ulery, Hicklin, Buscaglia, & Roberts, 2012).

The fact that experts reach different conclusions when examining the same data is an issue of *reliability*, which is distinct from the issue of *bias*, whereby exposure to irrelevant contextual information affects decision making. The Hierarchy of Expert Performance (HEP) presents research addressing reliability and bias as different and distinct elements in expert decision making (Dror, 2016).

Oliver's (2017) study reduced experts' disagreement (and increased confidence) not by increasing reliability in the way they draw conclusions, but by additionally providing all the experts with the same contextual information. Adding such contextual information can be appropriate, as long as

1. This contextual information is task-relevant. The National Commission on Forensic Science concluded that forensic experts should base their conclusions solely on relevant information (NCFS, 2015).
2. Examiners are transparent and explicitly state in their reports and testimonies that their conclusions are based on both the data and the contextual information.

Author Note
* Correspondence concerning this article should be addressed to Itiel E. Dror, University College London . Contact: i.dror@ucl.ac.uk

Oliver's commentary (2018) claims that Kukucka et al. (2017) misinterpreted his study by citing it as an example of how "irrelevant contextual information can distort [forensic examiners'] judgment" (p. 453). But the data clearly show that the contextual and historical information did in fact change the examiners' decisions in the Oliver (2017) study, as reflected in the increased consensus among them. However, this increased consensus is not necessarily an indicator of reliability or accuracy, but rather an artificial byproduct of giving all examiners the same irrelevant , non-medical contextual information. This finding hides the issue that examiners reach different conclusions when examining the same medical data (see HEP; Dror, 2016).

This raises a fundamental question about the appropriate role of the forensic expert. In our view, the forensic pathologist's role is to make decisions within his or her domain of expertise that are based solely on the medical-relevant data, not to integrate these data with other types of evidence. In the criminal and forensic context, the task of integrating multiple lines of evidence is appropriate for the detective, the District Attorney Office, the jury, and the judge—but not for the scientific expert (especially when they encounter evidence that is beyond their domain of expertise).

Our article cited Oliver (2017) as one of many "studies of professional forensic examiners. . . show[ing] that irrelevant contextual information *can* distort their judgment" (p. 453, emphasis added). The Oliver (2017) study clearly shows the impact of contextual information, and hence how it *can* distort judgment, which can increase error rates when the contextual information is misleading or incorrect. At the same time, we agree that relevant and accurate contextual information can be beneficial, and we clearly make that point in our Discussion, stating that "contextual information may be beneficial—or even essential—to an examiner's analysis" (p. 458).

We think and hope that we all agree that there are various factors that impact expert decision making, including that of medical examiners. These include issues of reliability, validity, bias, and effects of contextual information (both relevant and irrelevant to making a "medical" judgment). These factors need to be studied to illuminate whether and how they impact the decisions of experts.

Conflicts of Interest Statement

The authors declare no conflict of interest.

Author Contributions

All authors contributed.

References

Coble, M. (2015). *Interpretation errors detected in a NIST inter-laboratory study on DNA mixture interpretation in the U.S. (MIX13)*. Presentation at the international symposium on forensic science error management – Detection, measurement and mitigation, Washington, DC.

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

Dror, I. E., & Hampikian, G. (2011). Subjectivity and bias in forensic DNA mixture interpretation. *Science & Justice*, *51*(4), 204–208.

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.

Kukucka, J., Kassin, S. M., Zapf, P. A., & Dror, I. E. (2017). Cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *6*(4), 452–459.

National Commission on Forensic Science. (2015). *Ensuring that forensic analysis is based upon task-relevant information*. Available from https://www.justice.gov/ncfs/file/818196/download

Oliver, W. R. (2018). Response to Kukucka et al.: Comment on cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *7*(1), 158.

Oliver, W. R. (2017). Effects of history and context on forensic pathologist interpretation of photographs of patterned injury of the skin. *Journal of Forensic Sciences*, *62*(6), 1500–1505.

Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2012). Repeatability and reproducibility of decisions by latent fingerprint examiners. *PLoS ONE*, *7*(3), e32800.

Received 15 January 2018;
accepted 15 January 2018

Journal of Applied Research in Memory and Cognition 7 (2018) 314–315

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Comment on Dror, Kukucka, Kassin, and Zapf (2018), "When Expert Decision Making Goes Wrong"

William R. Oliver*

Regional Forensic Center, United States

Keywords: Cognitive bias, Forensic pathology, Diagnostic error, Forensic science

I am gratified by the authors' clear response. It explains a great deal about their advocacy for data hiding strategies to address what my community (forensic pathology) believes is a relatively minor problem compared to other more pressing issues. We both agree that "task irrelevant" information is not useful. However, to a physician, "task-relevant" information is information that is pertinent to achieving a correct diagnosis with respect to ground truth. To the authors, "task-relevant" information is information that is pertinent *and is in accordance with their model of the social role of the physician*. The authors state, "The expert's role is to make decisions based on the relevant data, not to integrate evidence. The task of integrating multiple lines of evidence is appropriate for the detective, the Distinct Attorney Office, the jury, and the judge—but not for the scientific expert." What we seem to disagree about is that social role.

What happens when that "evidence" *is* relevant data—as it so often is in medical diagnosis? That puts us pathologists in an awkward position. When we make a clinicopathologic correlation in the hospital setting, it is not merely acceptable for us to decide what is and is not "relevant" and to integrate multiple lines of evidence, it is mandatory. Part of our job is to discover pertinent contextual information that the clinician may have overlooked. It would be malpractice if we failed to do so. However, using the authors' model, making the same diagnosis in the same way with the same choices and the same inferential methods becomes "cognitive bias" if that case goes to court. One might further ask who should perform that integration in the 85–90% of cases we do that are primarily public health cases and never go to court. There is no District Attorney or jury of fry cooks and bus drivers to guide us through the nuances of integrating evidence of environmental exposure in the differential diagnosis of interstitial lung disease.

"Reliability" as used by the authors seems defined not as arriving at the correct diagnosis, but as arriving at the correct diagnosis *when the data is limited in accordance to their social model*. In my study, accuracy and reliability (in the traditional sense of the word) increased greatly by the addition of contextual information because that information was "relevant" to the diagnosis in the medical sense of the term. To physicians, this increase in accuracy does not "hide" (in the authors' words) unreliability, it is a component of the reliability that we achieve. Worse, proposals such as "sequential linear unmasking" seem more damaging than helpful if one is concerned with a medically correct diagnosis as opposed to a socially correct one, though that is a slightly different conversation (Oliver, Fudenberg, Howe, & Thomas, 2015).

These role-based distinctions explain much about why forensic pathologists do not consider this particular kind of cognitive bias as serious a problem as do the authors (though we are very concerned with other kinds of cognitive bias). It is always dangerous to do back-of-the-envelope calculations in an academic journal, but I will do it for illustration. Most large Medical Examiner offices in the United States have internal peer review systems to look for problems in diagnosis by individual pathologists. Errors are common, but usually do not involve diagnostic inference as much as process issues (copy editing, data entry, tissue processing, labeling, etc.). In my experience in four large systems, the rate of diagnostic disagreement due to inferential issues ran a little under one percent. A quick conversation with some colleagues in other offices resulted in similar results. So, while

Author Note

* Correspondence concerning this article should be addressed to William R. Oliver, Regional Forensic Center, 2761 Sullins St, Knoxville, TN 37919, United States. Contact: oliver@billoblog.com.

that is not a scientific survey, it is a reasonable guess, and is similar to more formal studies in surgical pathology where this subset of overall error runs between less than one to about five percent depending on case mix.

Of this one percent, most involve problems with knowledge base, nosology, and *failure* to incorporate pertinent contextual information. So, let's say of that 1% of cases, 20% are due to some sort of "cognitive bias" in the broad sense of the term, and of these, 5% are due to integration of "irrelevant" information in the medical (not social-role) sense of the term. Thus, I should run across a case where the peer review group will say, "Hey, you should have ignored this because it led you astray" about once every 10,000 cases, or once every 40 years of practice. That seems about right. In 32 years of practice, I can think of only one case where my diagnosis of cause of death was seriously challenged in court that way.

The authors may disagree with this calculation, but I doubt they have hard quantitative data to prove it wrong for forensic pathology as practiced in the United States. It is thus not surprising that I and my colleagues do not perceive an epidemic of misdiagnoses coming from the kind of cognitive bias that can be ameliorated by data hiding, nor do we see our diagnoses as somehow "artificially" correct. This is because there is no real evidence of that epidemic.

The authors label this as a "bias blind spot," but I do not think it really is. It is a disagreement about the role of the physician. Certainly, funding studies about error in forensic pathology comparable to those done in surgical pathology would be worthwhile. But there is simply no support for a claim of thousands of forensic pathologic misdiagnoses due to "cognitive bias" involving integration of data that is not "task-relevant," at least in the medical sense. Discussions about social roles can be fun. I do it myself (Oliver, 2014). But they are political and ideological, and have little to do with science, cognitive or otherwise (Dror, 2018).

Author Contribution

Dr. Oliver performed all functions.

Conflict of Interest Statement

The author declares no conflict of interest.

References

Dror, I. E. (2018). When expert decision making goes wrong: consensus, bias, the role of experts, and accuracy. *Journal of Applied Research in Memory and Cognition*, *7*(2018), 162–163.

Oliver, W. R. (2014). Manner determination in forensic pathology. *Academic Forensic Pathology*, *4*(4), 480–491.

Oliver, W. R., Fudenberg, J., Howe, J. A., & Thomas, L. C. (2015). Cognitive bias in medicolegal death investigation. *Academic Forensic Pathology*, *5*(4), 548–560.

Received 22 January 2018;
received in revised form 28 January 2018;
accepted 29 January 2018

Journal of Applied Research in Memory and Cognition 7 (2018) 316–317

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

No One is Immune to Contextual Bias—Not Even Forensic Pathologists

Itiel E. Dror*
University College London (UCL), United Kingdom

Jeff Kukucka
Towson University, United States

Saul M. Kassin and Patricia A. Zapf
John Jay College of Criminal Justice, United States

Our suggestions to improve forensic decision making involve well-established and accepted scientific methodologies.

1. Our approach aims to ensure that forensic pathologists base their decisions on relevant medical data while remaining blind to eyewitness identifications, confessions, and other non-medical evidence (which they are not trained to evaluate). As such, our approach is similar to the use of double-blind line-up identifications, the use of placebo controls in medical testing, the assessment of inter-rater reliability in psychological testing, and other strategies commonly used in science to minimize bias. These are not "data hiding strategies" (Oliver, in press) but, rather, well-established and universally accepted practices in applied science.
2. For information that is task-relevant, our strategies provide for its presentation to experts, but in a sequence aimed to minimize bias (such as suspect/target driven bias). Linear sequential unmasking (LSU; see Dror et al., 2015) provides the data, while optimizing what information is given, in what sequence, and when. LSU is a context management tool designed to minimize bias by linearly unmasking all the relevant information in the right sequence.

Although Oliver (in press) attributes the current disagreement to our "model of the social role of the physician," our objections are purely scientific. Ideally, forensic pathologists should make determinations based on relevant medical information, the interpretation of which falls within their training and expertise. For example, whether a person died as a result of suicide should be examined by the forensic pathologist based on characteristics of the wounds themselves, rather than by assessing whether the suicide letter is genuine or faked.

Yet Oliver argues that the pathologist should also consider a virtually unlimited range of non-medical data. Our view is that the pathologist should not be exposed to or consider such non-medical types of information. Forensic pathologists are not in a position, nor is it their job, to evaluate the accuracy of an eyewitness account, the voluntariness and credibility of a confession, or other lines of evidence not within their expertise.

The forensic pathologist is not an all-purpose expert looking to solve crimes by emulating "Quincy, M.E." (or its newer successor "Silent Witness"). The medical examiner's job description should be to provide a circumscribed and independent judgment of relevant medical evidence. It is then up to investigators, prosecutors, judges, and juries to integrate that medical judgment with other lines of evidence.

In Oliver's (2017) study, experts were given second-hand, non-medical information (e.g., "The girlfriend was seen by neighbors running from the house with what appeared to be a knife," p. 1503). This type of information did increase consensus and confidence, but the questions that need to be answered are why it did so, what it means, and is it warranted and appropriate. On the basis of this result, Oliver (2018) claims that contextual information "improved diagnoses by approximately 20% and

Author Note
* Correspondence concerning this article should be addressed to Itiel E. Dror, University College London, United Kingdom. Contact: i.dror@ucl.ac.uk

significantly decreased error." Yet it is impossible to make this claim given, as Oliver (2017) himself conceded, that "it was not possible to independently determine the 'ground truth' of the diagnosis" (p. 1501). It is therefore bewildering that he now claims that his study showed "decreased error" (Oliver, 2018) when his own published paper correctly describes it as "a study of consensus of diagnosis rather than correctness" (Oliver, 2017, p. 1501).

Even if contextual information leads all examiners to reach the same conclusion, this does not mean that their conclusion is accurate. The increased consensus and confidence can be based on evidence that should not be used, such as an eyewitness account that is wholly incorrect, a false confession, or information that is later ruled inadmissible.

To be sure, there are instances in which experts may need to integrate different forms of evidence. In those cases, scientifically sound protocols—such as the case manager approach (Dror, 2014) or linear sequential unmasking (LSU; Dror et al., 2015)—can be used to manage contextual influences in ways that minimize the risk of bias and cross-contamination (see *bias cascade* and *bias snowball* effects; Dror, Morgan, Rando, & Nakhaeizadeh, 2017).

Everyone's objective should be to minimize bias – not deny its existence. The Kukucka, Kassin, Zapf, and Dror (2017) survey clearly reveals a bias blind spot within the forensic sciences. Oliver's response just further illustrates this precise phenomenon.

Oliver's comments (in press) also underscore the importance and the need for data and scientific research—not introspection. The assertion that in 32 years of practice, I can think of only one case where my diagnosis of cause of death was seriously challenged in court" does not prove otherwise. One cannot infer that an expert was correct or mistaken based on whether his or her opinion was challenged in court. This point is well illustrated by the fact that nearly half of all exoneration cases in the U.S. involved people wrongfully convicted in part because of forensic science errors presented in court, often without challenge (Garrett & Neufeld, 2009).

Cognitive biases have been demonstrated in all realms of human decision making, including by experts, which is why it is important to use double-blind procedures, LSU, case managers and other safeguards that protect against bias. No one is immune—not even forensic pathologists.

Author Contributions

All authors contributed.

Conflicts of Interest Statement

The authors declare no conflict of interest.

References

Dror, I. E. (2014). Practical solutions to cognitive and human factor challenges in forensic science. *Forensic Science Policy & Management*, *4*, 105–113.

Dror, I. E., Morgan, R., Rando, C., & Nakhaeizadeh, S. (2017). The bias snowball and the bias cascade effects: Two distinct biases that may impact forensic decision making. *Journal of Forensic Sciences*, *62*(3), 832–833.

Dror, I. E., Thompson, W. C., Meissner, C. A., Kornfield, I., Krane, D., Saks, M., et al. (2015). Context management toolbox: A Linear Sequential Unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. *Journal of Forensic Sciences*, *60*(4), 1111–1112.

Garrett, B. L., & Neufeld, P. J. (2009). Invalid forensic science testimony and wrongful convictions. *Virginia Law Review*, *95*(1), 1–97.

Kukucka, J., Kassin, S. M., Zapf, P. A., & Dror, I. E. (2017). Cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *6*(4), 452–459. http://dx.doi.org/10.1016/j.jarmac.2017.09.001

Oliver, W. R. (2017). Effects of history and context on forensic pathologist interpretation of photographs of patterned injury of the skin. *Journal of Forensic Sciences*, http://dx.doi.org/10.1111/1556-4029.13449

Oliver, W. R. (2018). Response to Kukucka et al.: Comment on cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *7*(1), 161.

Oliver, W. R. (in press). Response to Dror et al.: Comment on when expert decision making goes wrong: Consensus, bias, the role of experts and accuracy. *Journal of Applied Research in Memory and Cognition, 7*(2).

Received 12 March 2018;
accepted 12 March 2018
Available online 16 April 2018