Doing Right by the Eyewitness Evidence: A Response to Berkowitz et al.

John T. Wixted[1], Laura Mickes[2], Chris R. Brewin[3], & Bernice Andrews[4]

[1]University of California, San Diego, [2]University of Bristol,[3]University College London, [4]Royal Holloway, University of London

John T. Wixted, Department of Psychology, University of California, San Diego. Laura Mickes, Department of Psychology, Royal Holloway, University of London.

Correspondence concerning this article should be addressed to John T. Wixted (jwixted@ucsd.edu).

Abstract

Berkowitz et al. (2020) attribute to us the claim that "confidence trumps all," and the few out-of-context quotations they selected can certainly be used to create that false impression. However, it is easily disproved, and we do so here. The notion that "confidence trumps all" is the mistake that the jurors made in the DNA exoneration cases, not a position that we have ever advocated.

Doing Right by the Eyewitness Evidence: A Response to Berkowitz et al.

According to Berkowitz et al. (2020), we claim that "confidence trumps all." To prove it, they quoted Brewin et al. (2020, p. 122) stating that "confidence is undeniably diagnostic of accuracy" in both laboratory studies and real criminal cases. However, they omitted context that disproves their claim. The actual argument was that confidence is undeniably diagnostic of accuracy "when fair lineups were created" and "solely during the first and only lineup test because confidence can increase over repeated testing" (see Wixted et al., 2016, p. 309, for a more detailed list of cautions about eyewitness confidence).

Berkowitz et al. (2020) also quoted a single statement from two days of Wixted's testimony in *People of the State of New York v. Boone* (2019), which they claim informed the court that "confidence trumps all." Once again, however, they inexplicably omitted context that refutes their claim. For example, he testified that high-confidence IDs are reliable (a) only when memory is not contaminated ("The first test is the uncontaminated test," p. 408), (b) only when the lineup is fair ("It's important that it not be an unfair lineup," pp. 405-406), (c) only when the suspect is identified quickly ("…seconds, not minutes," pp. 398-399), and (d) only in the absence of undue influence by the lineup administrator ("…if the cop says something like 'I need you to be sure,' that's not good," p. 417). The idea that "confidence trumps all" is the mistake that jurors made in the DNA exoneration cases, not a claim that we have ever made. Instead, time after time, we have explicitly enumerated factors that trump confidence.

Berkowitz et al. (2020) further allege that we have called on the field to "rethink the scientific nature of memory itself" because we believe that "memory is reliable". We instead wrote that memory "is clearly malleable but not unreliable under normal circumstances and in the absence of contamination or prolonged suggestion…" (Brewin et al., 2020, p. 4). Overall,

ours is a balanced perspective, one that cannot be fairly equated with the unnuanced perspective they incorrectly attribute to us.

That said, we agree with Berkowitz et al. (2020) that it is possible to oversimplify when condensing a complex argument into a single sentence—one that might later be quoted out of context. Everyone oversimplifies from time to time, including them. For example, Garrett (2011) wrote: "Although the most crucial and accurate information is how certain the eyewitness was at the time of the first identification, at that time most of these eyewitnesses were unsure" (p. 65). This sentence omits certain details (e.g., the eyewitnesses sometimes picked a filler or made no ID)—details they scold us for similarly omitting. Nevertheless, its message is essentially correct, and we are surprised that they now categorically disavow it, stating "…we do not have any idea how confident these eyewitnesses were in their initial identifications because there is no record" (p. 3). Of course there is a record. Otherwise, Garrett (2011) would have had no basis for claiming that most of the eyewitnesses were initially unsure. The record consists of trial testimony, including eyewitnesses themselves admitting under oath that, despite being confident now, they did not initially identify the suspect with high confidence—and sometimes did not identify the suspect at all.

Are their recollections believable? We think so. After all, witnesses are biased to recall their initial low-confidence IDs as having been made with high confidence (e. g. Wells & Bradfield, 1998), not the other way around. Moreover, people tend to recall the past in a way that is consistent with their current beliefs (e.g., Hoffrage et al., 2000). We know of no research suggesting that witnesses who are highly confident now—and who may be under pressure from a prosecutor to tell a consistent and compelling story—often falsely recall their past IDs as having been inconclusive. In our view, Garrett's (2011) results likely mean what he himself once

interpreted them to mean: that in most of the 161 cases, the initial memory test did not involve a high-confidence misidentification. Conceivably, very few did, and multiple independent lines of inquiry converge on that interpretation.

First, documented cases of initial misidentifications made quickly and with high confidence are hard to find. Second, a large body of lab-based research is now understood to show that suspect IDs made with high confidence are reliable (e.g., ~95% correct), contrary to what was long understood to be true. Third, although not addressed by Berkowitz et al. (2020), the same has been found to be true of eyewitnesses in the real world.

Wixted et al. (2016) reported results from 348 photo lineups administered to actual eyewitnesses by the Houston Police Department (PD). The suspects were strangers, the lineups were fair, and the lineup administrators were blind to the suspect's identity. Filler IDs (i.e., IDs to known-innocents) were made in 104 out of 348 lineups, which sounds like a lot. However, 84% were made with low or medium confidence. Thus, if we knew confidence for the initial misidentifications in the DNA exoneration cases (i.e., for those known-innocents), we might find that most were not made with high confidence. Garrett's (2011) results are consistent with that possibility despite real-world eyewitness identification procedures rarely being pristine.

High-confidence IDs were made in 88 out of the 348 Houston PD lineups. Remarkably, 81% of these IDs landed on suspects. Thus, despite being outnumbered by fillers 5 to 1, it seems that for IDs made with high confidence, the suspect often differentially matched the witness's memory of the perpetrator. One might speculate that these high-confidence suspect IDs reflect contaminated memory, but the available evidence weighs against that interpretation: when asked after the lineup test, witnesses rarely reported having seen the identified suspect other than at the crime scene. This makes sense. Except in rare cases (e.g., when the innocent suspect's face is

publicized or somehow found on social media), a witness would be unlikely to encounter the exact same innocent suspect that the police manage to find and place in a lineup.

Our case becomes even more compelling when we consider what theoretically should be true according to the standard framework that has guided thinking about recognition memory for decades (Wixted, 2020). Signal detection theory predicts that recognition memory errors made with low confidence should far outweigh errors made with high confidence. Thus, everything points in the same direction, and Berkowitz et al. (2020) present no confidence-related data or formal theoretical considerations to the contrary.

We agree with Berkowitz et al. (2020) that, despite the scarcity of documented cases, we cannot know for a fact that initial high-confidence misidentifications are uncommon in the real world. However, we can say with high confidence that multiple independent lines of theory and research point in that direction. Indeed, an inconclusive outcome on the initial test may be the norm in cases of misidentification, a fascinating possibility that has been completely overlooked by the field until recently. Because the very act of testing memory of a suspect irretrievably contaminates the witness's memory, police and prosecutors make a serious mistake when they ignore an initial test yielding an inconclusive (and presumptively unreliable) low-confidence ID. Under *those* conditions, confidence does trump all, and you can quote us on that.

References

Berkowitz, S. R., Garrett, B. L., Fenn, K. M., & Loftus, E. F. (2020). Convicting with confidence?

Why we should not over-rely on eyewitness confidence. *Memory*.

https://doi.org/10.1080/09658211.2020.1849308

Brewin, C. R., Andrews, B., & Mickes, L. (2020). Regaining consensus on

the reliability of memory. *Current Directions in Psychological Science, 29*, 121–125.

https://doi.org/10.1177/0963721419898122

Diamond, N. B., Armson, M. J., & Levine, B. (2020). The truth is out there: Accuracy in recall

of verifiable real-world events. *Psychological Science, 31*, 1544-1556.

https://doi.org/10.1177/0956797620954812

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Harvard

University Press.

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge

updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 566–581.

https://doi.org/10.1037/0278-7393.26.3.566

People of the State of New York v. Boone, 2190-11. (2019).

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect:" Feedback to

eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology,*

83, 360-376. https://doi.org/10.1037/0021-9010.83.3.360

Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 46, 201-233.

https://doi.org/10.1037/xlm0000732

Wixted, J. T., Mickes, L., Dunn, J. C., Clark, S. E., & Wells, W. (2016). Estimating the

reliability of eyewitness identifications from police lineups. *Proceedings of the National*

Academy of Sciences, 113, 304-309. https://doi.org/10.1073/pnas.1516814112