Reasons to Doubt the Reliability of Eyewitness Memory: Commentary on Wixted, Mickes, and

Fisher (2018)

Kimberley A. Wade[1], Robert A. Nash[2], D. Stephen Lindsay[3]

[1]University of Warwick

[2]Aston University

[3]University of Victoria

Author Note

[1] Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom

[2] Department of Psychology, Aston University, Birmingham, B4 7ET, United Kingdom

[3] Department of Psychology, University of Victoria, Victoria, BC, V8W 2Y2, Canada

Correspondence concerning this article should be addressed to Kimberley Wade, Department of

Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom. Email:

k.a.wade@warwick.ac.uk

Abstract

Wixted, Mickes, and Fisher (2018) take issue with the common trope that eyewitness memory is inherently unreliable. They draw on a large body of mock-crime research and a small number of field studies, which indicate that high-confidence eyewitness reports are usually accurate, at least when memory is uncontaminated and suitable interviewing procedures are used. We agree with the thrust of Wixted et al.'s argument, and welcome their invitation to confront the mass underselling of eyewitnesses' potential reliability. Nevertheless, we argue there is a comparable risk of overselling eyewitnesses' reliability. Wixted et al.'s reasoning implies that near-pristine conditions or uncontaminated memories are normative, but there are at least two good reasons to doubt this. First, psychological science does not yet offer a good understanding of how often and when eyewitness interviews might deviate from best-practice in ways that compromise the accuracy of witnesses' reports. Second, witnesses may frequently be exposed to pre-interview influences that could corrupt reports obtained in best-practice interviews.

Reasons to Doubt the Reliability of Eyewitness Memory: Commentary on Wixted, Mickes, and

Fisher (2018)

Wixted, Mickes, and Fisher (2018) challenge the common trope that eyewitness memory

is inherently unreliable (here, we use *reliable/unreliable* in the everyday sense, i.e., to be

dependable/undependable). Recollections are indeed malleable, they acknowledge, and can be

readily contaminated by improper interviewing techniques. But Wixted et al. assert that these

facts should not lead us to characterize eyewitness reports as an unreliable form of forensic

evidence. To support this assertion, they draw on mock-crime research and a small number of

field studies, which indicate that high-confidence eyewitness reports are usually accurate, at

least when memory is uncontaminated and suitable interviewing procedures are used. We

agree with the thrust of Wixted et al.'s argument, and welcome their invitation to confront the

mass underselling of eyewitnesses' potential reliability. Nevertheless, there is a comparable risk

of overselling eyewitnesses' reliability.

As we see it, Wixted et al.'s (2018) position is wholly compatible with the more common

characterization of eyewitness memory as *unreliable*. The compatibility of these positions hinges

on the base-rate of "pristine conditions" (or at least near-pristine conditions wherein initial

memory is uncompromised; Mickes, Clark, & Gronlund, 2017) in real-world eyewitness

interviews. If near-pristine conditions are rare, then both the common trope *and* Wixted et al.'s

characterization are plausibly correct.

Wixted et al.'s (2018) reasoning implies that near-pristine conditions or uncontaminated

memories are normative, but we doubt this. To consider why, we must look to psychological

science for answers to two questions. First, how often do eyewitness interviews deviate from best-practice and compromise the accuracy of witnesses' reports? Second, how often have witnesses been exposed to pre-interview influences that could corrupt reports obtained in best-practice interviews (cf. Poole & Lindsay, 1998)?

Answering the first question requires a definition of "best-practice." Wixted and Wells (2017), and many others (e.g., Levi & Lindsay, 2001; Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998) have sought to define the optimal conditions for eliciting identification decisions. Yet current evidence indicates that, despite promising shifts toward more widespread adherence, real-world investigators aren't consistently following these best-practice guidelines. For instance, in a survey of over 600 US law enforcement agencies, Loftus and Greenspan (2017) found that some best-practice procedures are frequently met (e.g., instructions that minimize pressure on witnesses to make positive identifications), whereas others are met less than one-third of the time (e.g., the use of double-blind lineups, explicitly probing witnesses' confidence in their identifications).

Psychology is in a far weaker position to define the optimal conditions for eliciting eyewitness descriptions of people or events. To say this is not to ignore the substantial literature on best-practices in information-gathering interviews (e.g., Hope, Gabbert, & Fisher, 2011; Poole & Lamb, 1998). But we know far less about the confidence-accuracy relationship under these different interviewing practices, and less still about how well this relationship holds in real-world cases. Note also that the method used for calculating the confidence-accuracy relationship can affect the assessments made and the conclusions that are drawn.[1] Unlike in the context of lineup identifications, we also have little evidence on how and when confidence

should be measured when witnesses provide narrative descriptions. The practicality and effectiveness of assessing confidence on a detail-by-detail basis in real-world interviews is yet to be established.

Even if we could define and guarantee near-pristine interviewing conditions in the real world, intra-interview procedures are not the only possible sources of memory contamination. Wixted et al. (2018) note that "only the first memory test counts" (p. 12), but how often is the official, recorded report the first occasion on which the witness has talked about the event? We do not know the answer, but research shows that co-witnesses frequently discuss crime details prior to being interviewed (Eisen, Gabbert, Ying, & Williams, 2017; Skagerberg & Wright, 2008). Moreover, it seems likely that victims of sexual assault and domestic violence often talk to medical personnel, friends, or family before they talk to police (Starzynski, Ullman, Filipas, & Townsend, 2005). Such discussions can make it difficult to determine the source of crime-relevant details (e.g., Vredeveldt, Groen, Ampt, & Van Koppen, 2017), and also create opportunities for memory contamination.

Lab studies show that people who receive misinformation are sometimes more confident in their erroneous memories than in their veridical memories (Assefi & Garry, 2003; Takarangi, Parker, & Garry, 2006), and become increasingly confident in those erroneous memories as rehearsal makes them increasingly coherent and rich (Lindsay, Hagen, Read, Wade, & Garry, 2004). Moreover, individuals' personal goals (Sharman & Calacouris, 2010), social goals (Hellmann, Echterhoff, Kopietz, Niemeier, & Memon, 2011), preferences (Gordon, Franklin, & Beck, 2005), expectations and stereotypes (Marsh, Cook, & Hicks, 2006), and semantic representations (Reyna, 1995) — to name just a few factors — can sometimes shape

remembering even in the absence of misinformation. These influences typically occur without conscious awareness (Johnson, Hashtroudi, & Lindsay, 1993), and therefore justice professionals, jurors, and witnesses themselves can seldom know whether or to what extent such factors have contaminated a memory. In sum, even if eyewitness recognition and recall are reliable when uncontaminated, it is risky to depend on that knowledge without having reliable ways of detecting contamination.

When is reliable, reliable enough? Wixted et al. (2018) claim that recall memory is highly reliable, citing evidence that when eyewitnesses report details with high confidence they are highly likely to be correct. But is it appropriate to describe witnesses as "reliable" when around 10% of the information they report with "absolute certainty" in near-pristine conditions is incorrect (as in Odinot, Wolters, & van Koppen, 2009)? If witnesses in near-pristine conditions are prompted to only answer questions when they are sure, does it make sense to describe these witnesses as "reliable" when 9% of their answers are wrong (as in Evans & Fisher, 2011)? Wixted et al. are right to note that these witnesses were overwhelmingly accurate, and we agree that witnesses have been impugned by common claims that memory is unreliable. Moreover, many of the memory errors made by witnesses are likely to have little forensic relevance. But as Wixted and colleagues have acknowledged elsewhere, with error-rates of 10% or more, even high-confidence memory reports are "simply not a reliable enough indicator of truth to unilaterally adjudicate guilt or innocence" (Roediger, Wixted, & DeSoto, 2012, p. 113). Therefore to characterize eyewitness memory as inherently reliable could risk leaving judges, juries, and investigators with the unwarranted assumption that 100% confidence implies 100% accuracy. Indeed, one important difference between eyewitness memory and DNA is that lay decision-

makers are rarely left to assess the reliability of the latter for themselves (see Thompson &

Newman, 2015).

It certainly seems timely to take stock of the robust and growing evidence on the

confidence-accuracy relationship, and to avoid treating eyewitness memories as specious by

default. But it would, in our view, be unwise to give the impression that eyewitness reports are

inherently reliable.

References

Assefi, S. L., & Garry, M. (2003). Absolut® Memory Distortions: Alcohol Placebos Influence the Misinformation Effect, *Psychological Science, 14*, 77- 80. DOI: 10.1111/1467-9280.01422

Eisen, M. L., Gabbert, F., Ying, R., & Williams, J. (2017). "I think he had a tattoo on his neck": How co-witness discussions about a perpetrator's description can affect eyewitness identification decisions. *Journal of Applied Research in Memory and Cognition, 6*, 274-282. DOI: 10.1016/j.jarmac.2017.01.009

Evans, J. R., & Fisher, R. P. (2001). Eyewitness memory: Balancing the accuracy, precision and quantity of information through metacognitive monitoring and control. *Applied Cognitive Psychology, 25*, 501–508. DOI: 10.1002/acp.1722

Gordon, R., Franklin, N., & Beck, J. (2005). Wishful thinking and source monitoring. *Memory & Cognition, 33*, 418 – 429. DOI: 10.3758/BF03193060

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using Receiver Operating Characteristic Analysis. *Current Directions in Psychological Science, 23*, 3-10. DOI: 10.1177/0963721413498891

Hellmann, J. H., Echterhoff, G., Kopietz, R., Niemeier, S., & Memon, A. (2011). Talking about visually perceived events: Communication effects on eyewitness memory, *European Journal of Social Psychology, 41*, 658–671. DOI: 10.1002/ejsp.796

Hope, L., Gabbert, F., & Fisher, R. P. (2011). From laboratory to the street: Capturing witness memory using the Self-Administered Interview. *Legal and Criminological Psychology, 16*, 211-226. DOI: 10.1111/j.2044-8333.2011.02015.x

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin,*
114, 3–28.

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in
eyewitness identification: Comments on what can be inferred from the low confidence-
accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,*
1304–1316. DOI:10.1037/0278-7393.22.5.1304

Levi, A. M., & Lindsay, R. C. L. (2001). Lineup and photo spread procedures: Issues concerning
policy recommendations. *Psychology, Public Policy, and Law, 7*, 776-790.

Lindsay, D.S., Hagen, L., Read, J.D., Wade, K.A., & Garry, M. (2004). True photographs and
false memories. *Psychological Science 15,* 149 -154. DOI: 10.1111/j.0956-7976.2004.01503002.x

Loftus, E. F., & Greenspan, R. L. (2017). If I'm certain, is it true? Accuracy and confidence in
eyewitness memory. *Psychological Science in the Public Interest, 18,* 1-2. DOI:
10.1177/1529100617699241

Marsh, R., Cook, G., & Hicks, J. (2006). Gender and orientation stereotypes bias source-
monitoring attributions. *Memory, 14,* 148-160. DOI: 10.1080/09658210544000015

Mickes, L., Clark, S. E., & Gronlund, S. D. (2017). Distilling the confidence-accuracy message: A
comment on Wixted and Wells (2017). *Psychological Science in the Public Interest, 18,* 6-9.
DOI: 10.1177/1529100617699240

Odinot, G., Wolters, G., & van Koppen, P. J. (2009). Eyewitness memory of a supermarket
robbery: A case study of accuracy and confidence after 3 months, *Law and Human Behavior,*
33, 506-514. DOI: 40540288

Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals.* Washington, DC, US: American Psychological Association.

Poole, D. A., & Lindsay, D. S. (1998). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Journal of Applied and Preventative Psychology, 7,* 1-26. DOI: 10.1016/S0962-1849(98)80019-X

Reyna, V. (1995). Interference Effects in Memory and Reasoning: A Fuzzy-Trace Theory Analysis. In F. N. Dempster & C. J. Brainerd, (Eds.). *Interference and Inhibition in Cognition* (pp. 29-59). Academic Press, New York.

Roediger, H. L., Wixted, J. T., & DeSoto, K. A. (2012). The Curious Complexity between Confidence and Accuracy in Reports from Memory. In L. Nadel, & W. Sinnott-Armstrong, (Eds.). *Memory and Law* (pp. 84 - 118). New York: Oxford University Press.

Sharman, S. J. & Calcouris, S. (2010). Do People's Motives Influence Their Susceptibility to Imagination Inflation? *Experimental Psychology, 57,* 77-82. DOI: 10.1027/1618-3169/a000010

Skagerberg, E. M., & Wright, D. B. (2008). The prevalence of co-witnesses and co-witness discussions in real eyewitnesses, *Psychology, Crime & Law, 14,* 513-521, DOI: 10.1080/10683160801948980

Starzynski, L. L., Ullman, S. E., Filipas, H. H., & Townsend, S. M. (2005). Correlates of women's sexual assault disclosure to informal and formal support sources. *Violence and Victims, 20,* 417-432.

Takarangi, M. K. T., Parker, S., & Garry, M. (2006). Modernising the Misinformation Effect: The Development of a New Stimulus Set, *Applied Cognitive Psychology, 20,* 583-590. DOI: 10.1002/acp.1209

Thompson, W. C., & Newman, E. J. (2015). Lay understanding of forensic statistics: Evaluation of random match probabilities, likelihood ratios, and verbal equivalents. *Law and human behavior*, *39*, 332-349.

Vredeveldt, A., Groen, R. N., Ampt, J. E., & Van Koppen, P. J. (2017). When discussion between eyewitness helps memory. *Legal and Criminological Psychology, 22*, 242-259.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603-647. DOI: 10.1023/A:1025750605807

Wells, G. L., Yang, Y., & Smalarz, L. (2015). Eyewitness identification: Bayesian information gain, base-rate effect equivalency curves, and reasonable suspicion. *Law and Human Behavior, 39*, 99-122.

Wixted, J., Mickes, L., & Fisher, R. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science.*

Wixted, J. T., Read, J. D., & Lindsay, D. S. (2016). The effect of retention interval on the eyewitness identification confidence-accuracy relationship. *Journal of Applied Research in Memory and Cognition, 5,* 192–203. DOI: 10.1016/j.jarmac.2016.04.006

Wixted, J. T. & Wells, G. L. (2017). The Relationship between Eyewitness Confidence and Identification Accuracy: A New Synthesis. *Psychological Science in the Public Interest, 18*, 10-65. DOI: 10.1177/1529100616686966

Footnotes

1   For comments on analytical techniques for measuring the accuracy-confidence

relationship, see Gronlund, Wixted, and Mickes (2014); Wells, Yang, and Smalarz, (2015);

Wixted and Wells (2017); Wixted, Read, and Lindsay (2016); and Juslin, Olsson, and Winman

(1996).