

Rethinking the Reliability of Eyewitness Memory

Perspectives on Psychological Science
2018, Vol. 13(3) 324–335
© The Author(s) 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691617734878
www.psychologicalscience.org/PPS



John T. Wixted¹, Laura Mickes², and Ronald P. Fisher³

¹Department of Psychology, University of California, San Diego; ²Department of Psychology, Royal Holloway, University of London; and ³Department of Psychology, Florida International University

Abstract

Although certain pockets within the broad field of academic psychology have come to appreciate that eyewitness memory is more reliable than was once believed, the prevailing view, by far, is that eyewitness memory is unreliable—a blanket assessment that increasingly pervades the legal system. On the surface, this verdict seems unavoidable: Research convincingly shows that memory is malleable, and eyewitness misidentifications are known to have played a role in most of the DNA exonerations of the innocent. However, we argue here that, like DNA evidence and other kinds of scientifically validated forensic evidence, eyewitness memory is reliable if it is not contaminated and if proper testing procedures are used. This conclusion applies to eyewitness memory broadly conceived, whether the test involves recognition (from a police lineup) or recall (during a police interview). From this perspective, eyewitness memory has been wrongfully convicted of mistakes that are better construed as having been committed by other actors in the legal system, not by the eyewitnesses themselves. Eyewitnesses typically provide *reliable* evidence on an initial, uncontaminated memory test, and this is true even for most of the wrongful convictions that were later reversed by DNA evidence.

Keywords

eyewitness identification, confidence and accuracy, Cognitive Interview

In the view of many, if there is one fact that has been conclusively established by psychological science over the past 30 to 40 years, it is that eyewitness memory is unreliable. And in one important way, there is no doubt that it is. Beginning in the 1970s, Elizabeth Loftus discovered the once-surprising but now uncontroversial fact that memory is *malleable*. With surprising ease, participants in a memory experiment can be led to believe, for example, that they saw a stop sign when they actually saw a yield sign (Loftus, Miller, & Burns, 1978) or that they became lost in a shopping mall as a child when no such experience actually occurred (Loftus & Pickrell, 1995).

The unfortunate malleability of memory has had tragic consequences in the legal system. For example, during the 1980s, a moral panic over day-care sexual abuse was later attributed to the unintentional implantation of false memories in young children during suggestive interviews (Bruck & Ceci, 1995; Ceci & Bruck, 1993; Ceci, Loftus, Leichtman, & Bruck, 1994). Likewise,

during the repressed-memory epidemic in the 1990s, adult patients in psychotherapy sometimes recovered childhood memories of having been sexually abused by their parents. Incredibly, parents were occasionally charged and convicted of sexual abuse on the basis of evidence consisting of nothing more than the recent recovery of a long-repressed memory from childhood. Only later did it become clear that many of the apparently recovered memories were actually unintentionally implanted by psychotherapists as they repeatedly probed a patient's childhood memories using techniques such as “guided imagery” (Loftus, 2003; Loftus & Ketcham, 1994). Although those moral panics have largely subsided, the malleability of memory continues to plague the legal system: Events that occur during

Corresponding Author:

John T. Wixted, Department of Psychology, University of California, San Diego, La Jolla, CA 92093
E-mail: jwixted@ucsd.edu

ordinary criminal investigations can have the effect of contaminating the memory of eyewitnesses, who then end up misidentifying innocent suspects or reporting events that did not occur. Indeed, eyewitness misidentifications are known to have played a role in 70% of the 353 convictions that have been overturned on the basis of DNA evidence since 1989 (Innocence Project, 2017).

In light of psychological research demonstrating the malleability of memory, and in light of the tragic consequences this has had in the legal system, it is perhaps not surprising that psychological science has rendered a verdict that now appears in virtually every textbook that addresses the issue: Eyewitness memory is unreliable. The purpose of this article is to suggest that it is time for that verdict to change. Against the notion that eyewitness memory is unreliable, we propose the following alternative perspective:

As is also true of other kinds of scientifically validated forensic evidence, eyewitness memory is *reliable* when it is not contaminated and when proper testing procedures are used.

This perspective concerning the reliability of eyewitness memory conflicts with what most researchers appear to believe and with what virtually every textbook that addresses the matter has to say about it.

The Prevailing Verdict on Eyewitness Memory

Evidence supporting the idea that eyewitness memory is widely perceived to be inherently unreliable is abundant. First, a search of Google using the exact phrase “eyewitness memory is unreliable” yielded 2,250 hits.¹ By contrast, a search of the exact phrase “eyewitness memory is reliable” yielded only 2 hits. Second, the Wikipedia entry on “eyewitness identification” (2017) quotes the late U.S. Supreme Court Justice William J. Brennan Jr.: “At least since *United States v. Wade*, 388 U.S. 218 (1967), the Court has recognized the inherently suspect qualities of eyewitness identification evidence, and described the evidence as ‘notoriously unreliable.’” The rest of the entry is written as if to validate Justice Brennan’s decades-old impression of eyewitness memory. Third, many psychology textbooks convey the message that eyewitness memory is unreliable, as readers can easily confirm for themselves if there is an introductory psychology text (or, perhaps, a social psychology text or memory text) on a nearby bookshelf. The following paragraph from a freely available online psychology text seems representative and will likely sound familiar to most readers:

Psychological researchers who began programs in the 1970s, however, have consistently articulated concerns about the accuracy of eyewitness identification. Using various methodologies, such as filmed events and live staged crimes, eyewitness researchers have noted that mistaken identification rates can be surprisingly high and that eyewitnesses often express certainty when they mistakenly select someone from a lineup. Although their findings were quite compelling to the researchers themselves, it was not until the late 1990s that criminal justice personnel began taking the research seriously. (*Cognitive Psychology and Cognitive Neuroscience/Memory*, 2017)

Are there any conditions under which eyewitness memory is highly reliable? If so, this textbook does not say. In our experience, the most any textbook ever does—beyond listing the *many* ways in which eyewitness memory can go wrong, complete with illustrations of real-life tragedies attributed to eyewitness misidentification—is to briefly acknowledge that eyewitness memory is not always inaccurate. Indeed, nearly every textbook treatment of eyewitness memory that we have seen is written as if the author’s primary responsibility is to disabuse readers of the dangerous idea that eyewitness memory might be a reliable form of forensic evidence.

Although we have not surveyed every psychology textbook, we feel safe in suggesting that no textbook leaves a reader with the impression that eyewitness memory is reliable in the same way that DNA evidence and fingerprint evidence are reliable (National Research Council, 2009; M. B. Thompson, Tangen, & McCarthy, 2014) when the evidence is not contaminated and when proper testing procedures are used. Instead, most textbooks leave the impression that is shared by almost everyone in our field, which is that eyewitness memory is simply unreliable. This widespread impression is now enshrined in amicus briefs on the reliability of eyewitness memory that have been filed by the American Psychological Association (APA). These legal documents state that their conclusions about eyewitness memory are based on scientific research and that they enjoy almost unanimous support in the field. A recent amicus brief states:

Importantly, error rates can be high even among the most confident witnesses. Researchers have performed studies that track, in addition to identification accuracy, the subjects’ estimates of their confidence in their identifications. In one article reporting results from an empirical study, researchers found that among witnesses who made

positive identifications, as many as 40 percent were mistaken, yet they declared themselves to be 90 percent to 100 percent confident in the accuracy of their identifications. (APA, 2014, pp. 17–18)

Another APA amicus brief asserts that “although the unreliability of eyewitness identifications is well known in the scientific community and among many lawyers, it is not understood by lay juries” (APA, 2016, p. 9). Similar assertions can be found in a recent amicus brief filed by the Innocence Project (2013).

It is not hard to understand how the field of psychology arrived at its generally negative assessment of the reliability of eyewitness memory. At least until the 1970s, and to some extent still today, the legal system operated as if the testimony of a credible and confident eyewitness was essentially infallible. Experimental psychologists in general (and Elizabeth Loftus in particular) awakened the legal system to the fact that eyewitness memory is malleable and is therefore not immune to contamination. It was a groundbreaking development that inspired new recommendations about forensic interviews and eyewitness identification procedures (e.g., Newlin et al., 2015; Police and Criminal Evidence Act, 2011; Technical Working Group for Eyewitness Evidence, 1999; G. L. Wells et al., 1998). Despite these positive developments, we submit that the once surprising revelation about the malleability of eyewitness memory has led to a severe overcorrection such that the field now regards eyewitness memory not only as potentially unreliable but also as *inherently* unreliable. In our view, the evidence does not support this idea and instead clearly refutes it.

When Is Forensic Evidence Reliable? The DNA Analogy

Before addressing the issue of the reliability of eyewitness memory, it is important to consider two points about the reliability of forensic evidence. First, few would dispute the idea that, as a general rule, forensic evidence of any kind can be contaminated. Thus, the fact that eyewitness memory can be contaminated is not a distinguishing feature of that type of evidence. Even DNA evidence can be contaminated, either before it arrives at the laboratory or if improper testing procedures are used in the laboratory itself. Indeed, there are multiple examples of evidence becoming contaminated with the DNA of an innocent person, ultimately resulting in a wrongful conviction (W. C. Thompson, 2013). Nevertheless, such cases are rare because police investigators and forensic DNA scientists are well aware of the risk of contamination, and they take appropriate steps to avoid that problem. With regard to laboratory

protocols, a document issued by the U.S. Federal Bureau of Investigation (FBI; 2001), “FBI Quality Assurance Standards Audit for Forensic DNA Testing Laboratories,” spells out requirements for annual laboratory audits and semiannual proficiency testing for DNA analysts to ensure that an accredited DNA laboratory is in compliance with FBI standards. When the forensic evidence is not contaminated and proper testing protocols are followed, DNA evidence is extraordinarily reliable (National Research Council, 2009). Thus, the mere fact that forensic evidence can be easily contaminated is not an automatic indictment of its reliability.

The second important point about judging the reliability of forensic evidence concerns the interpretation of a test result when (a) the evidence was not contaminated and (b) the test was properly performed. Critically, even under these conditions, a test result can be inconclusive. Again consider DNA evidence. Does a properly conducted DNA test of uncontaminated evidence from a crime scene conclusively identify the perpetrator or conclusively exclude an innocent suspect every time? Of course not. The results of a DNA test are summarized in a graph known as an *electropherogram*, which displays a series of sharp peaks reflecting the amount of DNA detected at various locations (loci) on the chromosomal material (W. C. Thompson, Ford, Doom, Raymer, & Krane, 2003). In a pristine single-source DNA profile, the electropherogram will exhibit either one or two peaks (representing alleles) at each of 20 different loci.² If the peaks of an unknown DNA profile obtained from the crime-scene evidence match all of the peaks from the known DNA profile obtained from a suspect, then the odds that the unknown sample was deposited by another person (i.e., not by the suspect) are infinitesimally small. However, sometimes, the crime-scene DNA evidence is degraded such that only a partial DNA profile is obtained (i.e., peaks are evident for only some of the 20 loci). This can occur even though the crime-scene evidence was not contaminated and even though proper testing procedures were followed in the crime laboratory. In that case, the results (i.e., a partial match) might not conclusively implicate the suspect.

An important component of a DNA test result is that it includes an indication of how definitive the results are. For example, the test result might indicate that whereas a partial profile of an unknown individual is consistent with the full DNA profile of the known suspect, there is a 1 in 4 chance that it would also be consistent with the full DNA profile of a randomly selected individual from the population. Under those conditions, the DNA test result would not constitute strong evidence against the suspect. Other test results might put the odds at 1 in 100, 1 in 100,000, or 1 in

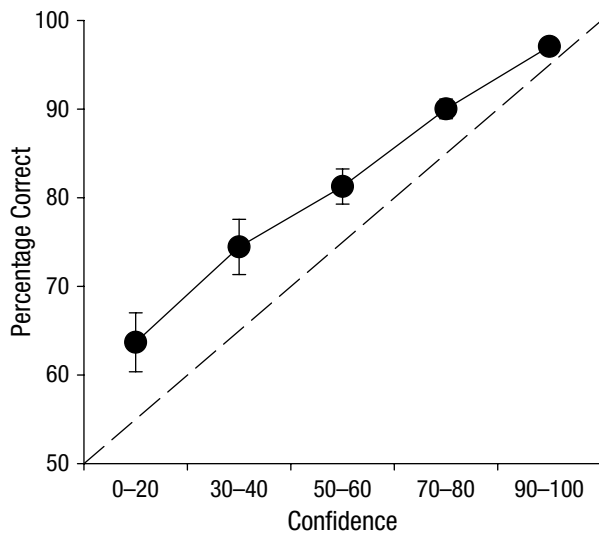


Fig. 1. Suspect-ID accuracy averaged across 15 studies with comparable scaling on the confidence (x-) axis. Error bars indicate ± 1 SD. Reprinted from Fig. 5a in Wixted and Wells (2017).

100 trillion, depending on the intactness of the DNA on the crime-scene evidence. The lower the odds that a randomly selected individual would yield a profile consistent with the DNA profile from the crime-scene evidence, the more certain one can be that the unknown DNA belongs to the suspect.

If it often happened that suspects were convicted as a result of a properly conducted DNA test result associated with a random-match probability of 1 in 4, many innocent people would end up in prison. However, as tragic as that would be, those wrongful convictions would not be an indication that DNA tests are inherently unreliable. Instead, it would be an indication that the criminal justice system is making a mistake by ignoring the random-match probability that accompanied the DNA test result. From this perspective, forensic evidence is reliable not because it provides accurate information whenever it is used in an effort to determine guilt or innocence (even DNA evidence is not reliable in that sense); rather, it is reliable if it includes a valid indication of the definitiveness of the evidence.

As with DNA, no account of the reliability of eyewitness memory is complete without considering the degree to which eyewitnesses can inform police investigators that the information they just provided is or is not definitive. The equivalent of the random-match probability in eyewitness memory is the confidence expressed by the eyewitness the first time memory is assessed. With that definition of reliability in mind, we next consider research pertaining to the reliability of eyewitness memory, first when it is tested by recognition (using a lineup) and then when it is tested by recall

(during a police interview). Although these two research literatures are usually considered separately, our argument will be that the same lesson has been learned in both cases: Eyewitness memory is reliable when the evidence is not contaminated, when proper procedures are followed, and when the confidence expressed by the eyewitness is taken into account.

Eyewitness Identification Evidence From a Lineup (Recognition)

Wixted, Mickes, Clark, Gronlund, and Roediger (2015) proposed that eyewitness identification evidence from a police lineup is highly reliable in the sense described above. That is, on an initial test of uncontaminated memory using proper procedures, low confidence implies low accuracy and high confidence implies high accuracy—as with DNA evidence. This is not to suggest that high-confidence eyewitness evidence can achieve the astronomically high levels of accuracy that can be achieved with DNA evidence (e.g., when the random-match probability is 1 in 100 trillion), but we do suggest that high-confidence IDs can achieve levels of accuracy that are far more impressive than is generally believed to be the case.

Recently, Wixted and Wells (2017) reviewed many laboratory studies and plotted the accuracy of a suspect identification (ID) as a function of confidence on a 100-point scale. The dependent measure used in their analyses addresses the question of greatest interest to judges and juries in a case involving eyewitness identification: Given that the witness identified the suspect with a certain level of confidence, what is the probability that the suspect ID was accurate? Figure 1 reproduces their summary figure, which was based on 15 simulated crime studies. Obviously, high confidence implies very high accuracy and low confidence implies much lower accuracy. Wixted, Mickes, Dunn, Clark, and Wells (2016) reported similar results from the Houston Police Department field study, shown here in Figure 2. The latter study is particularly important because it provided evidence that in actual practice (with real eyewitnesses), memory is reliable on an initial uncontaminated test using proper testing procedures.

Only the first memory test counts

The field of psychology has been slow to appreciate the strong relationship between confidence and accuracy—and, therefore, to appreciate the reliability of eyewitness identification—for at least three reasons. First, memory is malleable, and the very act of testing memory contaminates it by making it stronger than it

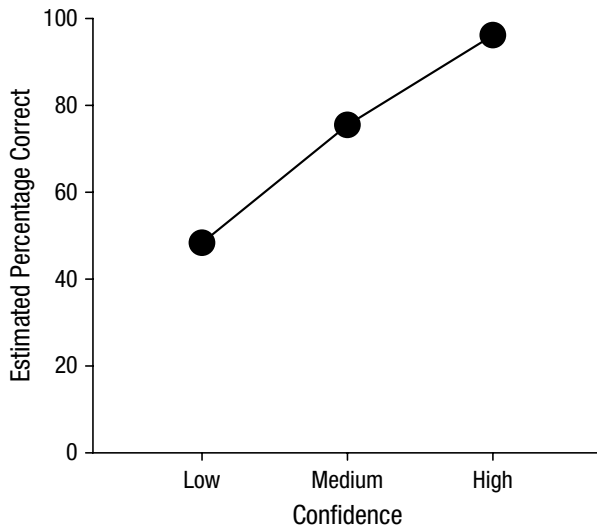


Fig. 2. Estimated suspect-ID accuracy as a function of confidence for the data from the Houston Police Department field study (based on Fig. 4b in Wixted, Mickes, Dunn, Clark, & Wells, 2016).

was before. Thus, on subsequent memory tests, eyewitnesses will experience a stronger memory-match signal than they did on the first test and more confidently identify a suspect—whether that suspect is innocent or guilty (e.g., Steblay & Dysart, 2016). The implication is that only the first test of an eyewitness’s memory can provide untainted forensic evidence (Wixted et al., 2015), and that fact needs to be considered in any discussion about the reliability of eyewitness memory. Many have conflated contaminated memory tests (i.e., the eyewitness identification test that occurs at trial) with uncontaminated memory tests (i.e., the initial memory test conducted using a properly constructed lineup) when forming an opinion about the reliability of eyewitness memory. Yes, contaminated memory evidence is unreliable (just as contaminated DNA evidence is unreliable), but no, that fact does not indicate that eyewitness memory is inherently unreliable. To ask whether eyewitness evidence (or DNA evidence or fingerprint evidence) is reliable is to ask whether it is reliable when the evidence is not contaminated, when proper testing procedures are used, and when confidence is taken into account. For eyewitness identification, because only the first test provides a test of uncontaminated memory, its reliability must be judged in relation to that first test.

The “correlation” between confidence and accuracy is irrelevant

The second reason why the field has been slow to appreciate that eyewitness identification evidence is reliable is that early investigations into the confidence-accuracy

relationship suggested that the relationship is weak or, at best, moderate, even on an initial uncontaminated test using a pristine lineup procedure (Penrod & Cutler, 1995; Sporer, Penrod, Read, & Cutler, 1995). If that were true, it would certainly support the claim that eyewitness identification evidence is inherently unreliable. However, it is now understood that the original conclusion about a weak-to-moderate confidence-accuracy relationship was based on the problematic use of a statistic—the point-biserial correlation coefficient—to measure that relationship (Juslin, Olsson, & Winman, 1996; Wixted & Wells, 2017). The most straightforward way to communicate the confidence-accuracy relationship for cases in which a suspect has been identified from a lineup is to simply plot suspect-ID accuracy as a function of confidence (Mickes, 2015), as we have done here in Figures 1 and 2. When plotted that way, the confidence-accuracy relationship is impressively strong.

Suboptimal memory conditions

The third reason why eyewitness-identification evidence has long been judged to be unreliable is the belief that for a strong confidence-accuracy relationship to hold, not only must an appropriate test be administered, but the memory conditions at the time of the crime must also be favorable. For example, the “optimality hypothesis” (Deffenbacher, 1980, 2008) holds that confidence becomes less indicative of accuracy under suboptimal conditions (e.g., high stress, the presence of a weapon, racial differences between the witness and the perpetrator, short exposure duration, long retention interval). However, the evidence collected to date suggests that variables such as these do not appreciably affect the accuracy of initial identifications made with high confidence (Mickes, 2015). The relevant studies were reviewed by Wixted and Wells (2017), and key results were summarized in their Figures 4c (weapon present or absent), 4d (weapon present or absent again), 4f (same race vs. cross race), 4h (short vs. long retention intervals), 4i (short vs. long retention intervals and short vs. long exposure durations), 4m (full vs. divided attention), 4n (short vs. long retention interval), and 4p (short vs. long retention interval). Although these variables had the expected negative effect on the overall accuracy of memory (because, for example, memory fades after a long retention interval), they did not diminish the accuracy of suspect IDs made with high confidence.

The results of these studies help to further clarify the difference between “reliability” and overall “accuracy.” To say that a long retention interval reduces the overall accuracy of eyewitness memory means that eyewitness will make more mistakes after a long retention

interval compared with a short retention interval. However, it does not automatically follow that the information provided by eyewitnesses when they make a suspect ID under unfavorable memory conditions is any less reliable. For example, after a short retention interval, almost all suspect IDs might be both accurate and accompanied by high confidence. After a long retention interval, some suspect IDs might be accurate and accompanied by high confidence, but most suspect IDs might be inaccurate and accompanied by low confidence. Overall suspect-ID accuracy will be lower after a long retention interval because most of what is remembered consists of low-confidence, low-accuracy IDs, but reliability might be unaffected in that high-confidence IDs, although less likely to occur, are as accurate as ever. It is therefore no contradiction to say that variables such as high stress, weapon focus, and long retention interval reduce overall accuracy without necessarily affecting the reliability of suspect IDs made with high confidence.

The results of these laboratory studies help to make sense of the police department field study results summarized here in Figure 2. In the Robbery Division of the Houston Police Department, more than 60% of the IDs were cross-race IDs, and more than 70% of robberies involved the presence of a weapon (W. Wells, Campbell, Li, & Swindle, 2016). Given that a weapon was often present, one can reasonably assume that eyewitness stress is often high. Even so, just as in laboratory studies involving other variables that negatively affect overall accuracy, high-confidence IDs were estimated to be highly accurate.

With these findings in mind, it is worth revisiting the DNA exoneration cases—70% of which were based on eyewitness misidentifications. In his book *Convicting the Innocent: Where Criminal Prosecutions Go Wrong*, Garrett (2011) analyzed trial materials for 161 DNA exonerees who had been misidentified by one or more eyewitnesses in a court of law. In every case, the eyewitness testified with high confidence at trial that the defendant was the perpetrator. However, because IDs that occur at trial are not initial IDs, that fact by itself shows only that contaminated memory is unreliable (something that was not widely appreciated in the pre-Loftus era but now is, at least among experimental psychologists). Critically, in 57% of those cases, information was available about the level of confidence expressed on the initial (presumably uncontaminated) memory test. In every one of those cases, the same eyewitnesses who were highly confident in their misidentifications at trial were initially *uncertain*, at best. In other words, on the one and only test that counts (the initial uncontaminated memory test), the result was a lot like an inconclusive DNA test that comes back

indicating that the DNA profile on the crime-scene evidence, although consistent with the suspect's DNA, is also consistent with DNA from a high proportion of the population. Presumably, most prosecutors would not try to convict someone on the strength of such weak evidence alone. The fact that similarly weak eyewitness evidence from an initial test ended up convicting a large number of innocent defendants does not mean that eyewitness memory is inherently unreliable. Instead, it means that the legal system ignored the results of a valid initial test that was based on uncontaminated eyewitness evidence and instead unwittingly used contaminated eyewitness evidence to win a conviction.

Eyewitness Evidence From a Police Interview (Recall)

The considerations discussed above pertain to eyewitness identification (recognition memory), but similar considerations apply to information obtained from interviewing eyewitnesses about a crime they observed (recall memory). For example, just as lineup administrators can distort eyewitness recognition to generate incorrect identifications, so too can interviewers distort eyewitness recall to generate incorrect descriptions. Such error-inducing techniques include (a) asking suggestive questions or otherwise introducing postevent misinformation (Loftus et al., 1978), (b) asking an abundance of closed questions (vs. open-ended questions; Lamb, Orbach, Hershkowitz, Horowitz, & Abbott, 2007), and (c) encouraging or enticing witnesses to guess (vs. providing an option not to respond; Earhart, La Rooy, Brubacher, & Lamb, 2014). When these avoidable error-inducing interview techniques are avoided, however, as in the cognitive interview (Fisher & Geiselman, 1992) or the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) protocol (Orbach et al., 2000), witness descriptions can be quite accurate. We next consider research on the accuracy of information that is elicited from adults using proper interview techniques in laboratory studies and police-department field studies.

Laboratory studies of eyewitness recall

Laboratory studies typically have participants view a mock crime (as in laboratory studies of eyewitness identification), after which different interviewing strategies are compared (e.g., the cognitive interview vs. a standard police interview protocol). Although most studies are concerned with the ability of a structured interview such as the cognitive interview to elicit *more* information than alternative techniques, many studies also report the accuracy of the obtained information.

In an early meta-analysis of the literature, Koehnken, Milne, Memon, and Bull (1999) determined that the cognitive interview elicited significantly more information than alternative interviews. More important for our present purposes, they also found that “accuracy (as measured by the proportion of all witness statements that were correct) was as high or slightly higher in the [cognitive] interviews (accuracy rate = 0.85) than in the comparison interviews (0.82)” (p. 63). Comparable accuracy scores were recently reported by Rivard, Fisher, Robertson, and Hirn Mueller (2014) in a more realistic study in which participants were asked to recall the details of a meeting they had attended at a training center for federal investigators 3 to 43 days earlier. The interviewers were staff members who conduct training programs on investigative interviewing at the training center. The cognitive interview again outperformed a standard interview in terms of the amount of information elicited (the usual result). Critically, of the recalled information that could be corroborated (based on records of the meeting provided by the training facility), the accuracy rates were 88% correct and 87% correct for the cognitive interview and a standard interview protocol, respectively. Accuracy rates that fall in the range of 85% to 90% correct indicate that eyewitness recall for details is not perfect, but such results are hard to reconcile with the prevailing notion that eyewitness memory is simply unreliable.

The high accuracy rates noted above, which were found in controlled laboratory studies, are instructive but may not be representative of performance in real-world investigations. Note that most laboratory studies (a) use undergraduate student witnesses, who may have better memory or verbal skills than typical victims and witnesses of real crime, and (b) use graduate research assistant interviewers, who may ask easier questions than real-world police investigators—although as noted above, Rivard et al. (2014) used nonstudent witnesses and professional interviewers (see also Mosser, Fisher, Molinaro, Satin, & Manon, 2016). Furthermore, laboratory studies are often designed to detect differences across conditions and so they are engineered to avoid ceiling and floor effects, in which case the absolute levels of performance may not be informative. We should therefore look to field studies of actual victims and witnesses to crime to see whether the high witness accuracy rates still hold.

Field studies of eyewitness recall

In an early police-department field study, Fisher, Geiselman, and Amador (1989) trained seven experienced detectives from the Robbery Division of the Metro-Dade Police Department to use the cognitive

interview and compared their performance against that of nine untrained but equally experienced detectives. As in laboratory studies, the trained detectives elicited considerably more information than the untrained detectives did. Because the ground truth of recalled information was unknown, it was estimated using various sources of corroboration (mainly reports from other witnesses interviewed immediately after the crime). Overall corroboration rates exceeded 93% for both the cognitive interview and the standard interview protocol. These findings suggest that, if anything, real eyewitnesses may be slightly *more* accurate than what is suggested by laboratory studies (see also van Koppen & Lochun, 1993; Yuille & Cutshall, 1986). One obvious limitation of this interpretation is that the measure of accuracy used in this study—corroboration across witnesses—is imperfect, because it is possible for two witnesses to be consistent with each other but to still be wrong. More recent studies of real-crime witnesses relied on purer measures of accuracy.

Several studies have taken advantage of the fact that crimes are sometimes captured on closed-circuit television (CCTV). Those CCTV images can then be used to directly validate the information later recalled by eyewitnesses. In the first study making use of CCTV images, Woolnough and MacLeod (2001) examined archived police records and identified eight incidents of assault that involved both a victim and a bystander. Most of the elicited information was about the crime events, not descriptions of the perpetrator. The action details recalled by the witnesses included events such as the victim being knocked to the ground, a man and a woman having an argument and pushing each other, good Samaritans breaking up a fight, and so forth. In this study, victims and bystanders were found to be highly accurate: Both achieved accuracy scores of 96% correct. In another CCTV-corroborated field study, eyewitnesses provided descriptions of the perpetrators of armed robberies in Oslo, Norway (Fahsing, Ask, & Granhag, 2004). Of the verifiable attributes, 87% were correct. Again, findings such as these seem impossible to reconcile with the notion that eyewitness memory is generally unreliable.

Keep in mind that our focus is on an initial test of memory that occurs before experiences that may contaminate witness memory (just as DNA evidence can be contaminated). A good illustration of how contamination can reduce the reliability of information obtained from a police interview comes from an archival police study of 29 people who witnessed the murder of Swedish Foreign Minister Anna Lindh (Granhag, Ask, Rebelius, Öhman, & Giolla, 2013). In this case, only 58% of the reported attributes were correct, as corroborated by CCTV. According to the authors, the most likely

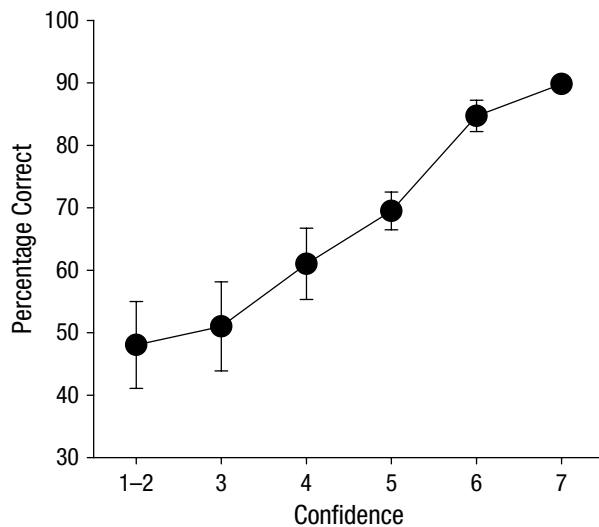


Fig. 3. Observed relationship between percentage of recalled details that were correct and confidence. Error bars indicate ± 1 SE. Data are from Roberts and Higham (2002).

explanation for the poor performance was memory contamination that occurred because the witnesses were gathered together before being interviewed, and they discussed the event. These findings underscore the fact that our claims about the surprisingly high reliability of eyewitness memory pertain to tests of memory that are conducted before memory contamination.

In all, laboratory studies of eyewitness memory that use generally accepted interviewing protocols and do not intentionally provide misleading information or entice witnesses to guess find that accuracy is quite high (~85%–90%). Field studies of police interviews with victims and witnesses of real crime show, if anything, even higher rates of accuracy. On the surface, this finding might seem hard to reconcile with reports showing that police interviewers often do not follow optimal interview procedures (Fisher, Geiselman, & Amador, 1989; Snook & Keating, 2011). However, analyses of police interviews show that the police are likely to elicit less information than is potentially available: They do not typically use techniques that increase the risk of eliciting inaccurate information, and they typically do not make use of empirically validated techniques for maximizing the quantity of recalled information, such as the techniques used in the cognitive interview (e.g., context reinstatement, witness-compatible questions, encourage active witness participation). Nevertheless, at least when interviewing cooperative witnesses, the police rarely ask blatantly suggestive questions or offer misleading information, procedures that are known to contribute to witness error. Thus, we would expect to find that training police to use the cognitive interview increases the amount of

information elicited but not the accuracy of witness reports because accuracy is already high.

Eyewitness confidence and recall accuracy

When discussing eyewitness identification (recognition memory) earlier, we noted that, under proper testing conditions, confidence was a good indicator of accuracy. How well does eyewitness confidence indicate accuracy when eyewitnesses are describing a recollected event (recall memory)? Several laboratory and field studies converge on the conclusion that eyewitness confidence also predicts recall accuracy.

Roberts and Higham (2002) conducted one of the first laboratory studies to examine confidence as a predictor of eyewitness recall accuracy. After watching a videotape of a simulated robbery, laboratory witnesses were interviewed about their recollections of the event. After an initial free narrative report, witnesses were asked follow-up questions to elaborate on their initial recollections. Witnesses were then asked to make confidence judgments (using a 1-to-7 scale) about each detail that was reported earlier. We estimated the overall number of correct and incorrect details for each level of confidence from their Figure 1 and then computed the probability that a recalled detail was correct for each level of confidence. That is, for each level of confidence, we calculated this probability as follows: number of correct details / (number of correct details + number of incorrect details). We used the result to create the recall version of a confidence-accuracy characteristic (CAC) plot. Because very few details were recollected with confidence ratings of 1 or 2, we collapsed across those two confidence ratings. The results are shown in Figure 3. Obviously, as with the recognition studies considered earlier, confidence was strongly related to accuracy, and high confidence was associated with high accuracy.

Odinot, Wolters, and van Koppen (2009) extended the Roberts and Higham (2002) laboratory study to a more naturalistic and stressful setting, an armed robbery of a supermarket. Three months after the crime, Odinot et al. interviewed eyewitnesses about details of the crime. After the witnesses described the event, the experimenters asked the witnesses to make confidence judgments (on a 1-to-7 scale) about each detail they had provided earlier. Odinot et al. (2009) measured the relationship between confidence in recollected details and the accuracy of those recollections by computing a γ correlation coefficient (which, in their study, came to .38), but that approach suffers from the same problem as the point-biserial correlation coefficient used in eyewitness identification studies: It is capable of masking a strong relationship (Roediger, Wixted, & DeSoto, 2012). Instead

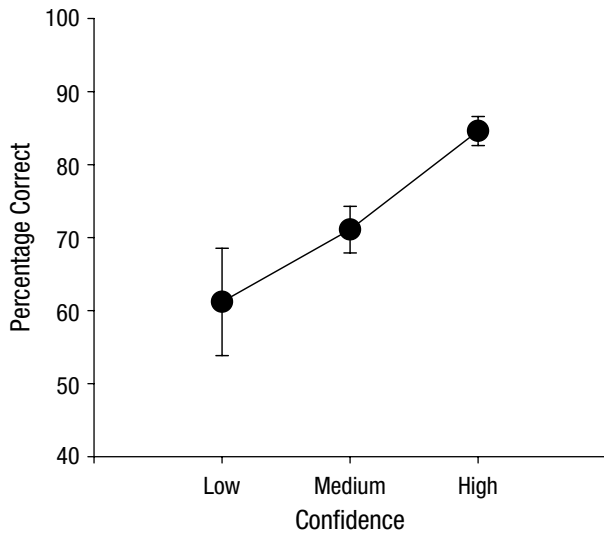


Fig. 4. Observed relationship between percentage of recalled details that were correct and confidence. Error bars indicate ± 1 SE. Data are from Odinet, Wolters, and van Koppen (2009).

of computing a correlation coefficient, it is more informative to simply plot the relationship between confidence and accuracy in a manner similar to the CAC plots shown here in Figures 1, 2, and 3. The information needed to do so was reported by Odinet et al. (2009) in their Table 2. To create the recall version of the CAC plot, we averaged across different categories (person descriptions, object descriptions, and action details) for the nine central witnesses interviewed in that study. In addition, because many responses were associated with high confidence (a rating of 7), whereas relatively few responses were made with confidence ratings of 1 through 6, we computed weighted averages across categories to create a 3-point scale (low, medium, and high confidence). Low confidence consisted of ratings of 1 through 3 (44 ratings in all), medium confidence consisted of ratings 4 through 6 (203 ratings), and high confidence consisted of ratings of 7 (326 ratings). Figure 4 shows that correct recall increased from 61% to 85% as confidence increased from low to high. Again, high-confidence recollections, at 85% correct, despite being less than perfect, cannot be characterized as being unreliable.

Even though these findings are encouraging, we suspect that they may underestimate the predictive power of confidence: In both the Roberts and Higham (2002) and Odinet et al. (2009) studies, the confidence judgments were made only after a delay (after witnesses reported several related details)—that is, these were not contemporaneous confidence judgments. Requiring interviewees to make a confidence judgment immediately after each recalled fact would disrupt the natural flow of a properly conducted interview. (We note that although it is very natural for witnesses doing a

recognition/identification test to be asked immediately after each decision to indicate their level of confidence—e.g., “You just picked number 5; how confident are you about that decision?”—it may be unnatural for a witness interview.) In order to examine confidence judgments made spontaneously within a recall task without disrupting the natural flow of a witness interview, other indirect measures of confidence can be used. Two such measures, which we discuss below, are (a) noticing when witnesses spontaneously use verbal expressions of certainty or uncertainty and (b) allowing witnesses to withhold low-confidence recollections.

Paulo, Albuquerque, and Bull (2016) showed experimental witnesses a videotape of a simulated crime and then interviewed the witnesses by eliciting mainly uninterrupted narrative descriptions of the event. Whereas most of the witness statements were unqualified (e.g., “it was a red shirt”), they sometimes uttered verbal expressions to convey their uncertainty (e.g., “I think it was a red shirt” or “maybe it was a red shirt”). In keeping with the general principle that certainty indicates heightened accuracy, unqualified statements were correct 90% of the time, whereas statements preceded by expressions of uncertainty were correct only 65% of the time. Thus, even though the interviewers did not formally ask witnesses to indicate their level of certainty, spontaneous measures of certainty and uncertainty still emerged to distinguish between highly accurate and less accurate witness recollections.

Finally, witness confidence may be assessed unobtrusively by allowing witnesses to withhold uncertain responses or to say “I don’t know” (Koriat & Goldsmith, 1994). Presumably, if given the opportunity to withhold a response, witnesses will withhold the low-confidence responses. In a test of this hypothesis, Evans and Fisher (2011) showed experimental witnesses a brief crime video and then tested them with either free recall, cued recall, or yes/no questions. Moreover, witnesses were either permitted to withhold responses (say “I don’t know”) or they were forced to answer all questions. As might be expected, witnesses were less confident about responses they withheld (and were later asked to report) than responses they provided voluntarily. In agreement with the results summarized in Figures 3 and 4 witnesses were very accurate for the high-confidence answers they provided voluntarily (probability of being correct = .91), but accuracy declined considerably when they were forced to also provide low-confidence responses (probability of being correct = .79).

Conclusion

Our main message is that when investigators probe eyewitness memory, either via identification procedures

(recognition tests) or interviews (recall tests), the information they receive is likely to be very reliable if the following conditions are met: (a) Witnesses were not previously exposed to distorting or contaminating information, (b) the witness' memory is being probed for the first time, (c) witnesses are not "tricked" into providing desired information (e.g., through the use of biased lineups or suggestive interview questions), (d) the witness's metacognitive monitoring guides his or her responding (either by withholding a response if uncertain or explicitly reporting his or her level of confidence), and (e) the investigator is sensitive to the witness's level of confidence (i.e., relying on high-confidence responses while attaching less weight to low-confidence responses). When such conditions are met, eyewitness memory is likely to be reliable—both in the laboratory and in the field—whether the test involves recognition or recall. When such conditions are not met, eyewitness memory may be unreliable—but that is hardly because of a faulty memory system.

These considerations indicate how the message from experimental psychology—namely, that eyewitness memory is inherently unreliable and that eyewitness confidence should be disregarded—is incomplete, to say the least. The evidence we have reviewed here indicates that eyewitness memory is reliable in the same way that DNA evidence is reliable. When proper procedures are used, both DNA test results and eyewitness-memory test results are accompanied by an indication of the information's reliability. For DNA tests, that information consists of the random match probability. For memory tests, that information consists of an eyewitness's confidence in the information that was just provided (with respect to either an ID made from a lineup or an answer provided in response to an interview question). Ignoring that critical piece of information can lead to tragic errors, including wrongful convictions of the innocent. Indeed, ignoring the low confidence expressed by eyewitnesses on the initial memory test is exactly how most of the innocent defendants who were supposedly wrongfully convicted because of the unreliability of eyewitness memory ended up in prison in the first place (only to be exonerated by DNA evidence years later). The same would be true if people were routinely convicted on the basis of inconclusive DNA evidence. Fortunately, so far as we know, the legal system does not make that mistake when the forensic evidence involves a DNA test (i.e., if the test is inconclusive, that evidence is not used to prosecute), but it does make that mistake when the forensic evidence consists of an eyewitness-memory test. Blaming the inevitable wrongful convictions on the unreliability of eyewitness memory is pointing the finger of blame in the wrong direction.

It might be argued that the perspective we have advanced here is defensible in theory, but that, in practice, eyewitness evidence is so often mishandled that it is nevertheless valid to assert that eyewitness memory is (for all practical purposes, anyway) inherently unreliable. However, for two reasons, we believe that this is not a viable position. First, keep in mind that in the DNA exoneration cases for which the nature of the initial ID could be determined, the witnesses did not express high confidence. In fact, in no such case was a witness both mistaken and highly confident (Garrett, 2011). These findings provide direct evidence that, by the time of the first ID in a typical police investigation, eyewitness memory is usually not contaminated to the point where a mistaken ID will happen with high confidence. That obviously can happen, but the available evidence suggests that it is not a frequent occurrence. If it were, one would expect to find many cases in which the initial ID in a DNA exoneration case were made with high confidence. So far, there is no such evidence.

Second, many jurisdictions have adopted much improved eyewitness memory protocols in recent years. A recent memo from the U.S. Department of Justice, for example, instructed all federal law enforcement agencies to adopt "best practices" eyewitness-identification protocols: A lineup should be fair (i.e., the suspect should not stand out), it should contain only one suspect, and that an initial statement of confidence should be obtained (Yates, 2017). In federal trials involving eyewitness-identification evidence, should juries be told that eyewitness memory is inherently unreliable even if the DOJ guidelines were followed? That seems inappropriate to us. Instead, just as is true of trials involving DNA evidence, the jury should hear arguments about whether proper testing protocols were adhered to so the jury can make an independent judgment about the reliability of the evidence. When memory is not contaminated and proper testing procedures are followed, eyewitness memory is clearly reliable. In our opinion, the cause of justice is not served by suggesting otherwise.

Acknowledgments

The content of this article is solely the responsibility of the authors and does not necessarily reflect the views of the National Science Foundation, the Economic and Social Research Council, or the U.S. Department of Justice.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported in part by National Science Foundation Grant SES-1456571 and Economic and Social Research Council Grant ES/L012642/1 (to L. Mickes and J. T. Wixted) and by U.S. Department of Justice, Federal Bureau of Investigation Contract Number DJF-15-1200-V-0010434 (to R. P. Fisher).

Notes

1. The searches were conducted on February 11, 2017. The phrases we searched for included the quotation marks.
2. In 1997, the FBI selected 13 loci to constitute the core of the United States Combined DNA Index System (CODIS), but as of January 1, 2017, the number of core loci was increased to 20.

References

- American Psychological Association. (2014). *Brief for Amicus Curiae, Commonwealth v. Gomes and Commonwealth v. Johnson*. Retrieved from <http://www.apa.org/about/offices/ogc/amicus/gomes-johnson.pdf>
- American Psychological Association. (2016). *Brief for Amicus Curiae, Payne v. Commonwealth of Virginia*. Retrieved from <http://www.apa.org/about/offices/ogc/amicus/payne.pdf>
- Bruck, M., & Ceci, S. J. (1995). Amicus brief for the case of State of New Jersey v. Michaels presented by Committee of Concerned Social Scientists. *Psychology, Public Policy, and Law, 1*, 272–322.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.
- Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 42*, 304–320.
- Cognitive psychology and cognitive neuroscience/memory. (2017, June 24). In Wikibooks. Retrieved October 10, 2017 from https://en.wikibooks.org/w/index.php?title=Cognitive_Psychology_and_Cognitive_Neuroscience/Memory&oldid=3237400
- Deffenbacher, K. A. (1980). Eyewitness accuracy and confidence: Can we infer anything about their relationship? *Law and Human Behavior, 4*, 243–260.
- Deffenbacher, K. A. (2008). Estimating the impact of estimator variables on eyewitness identification: A fruitful marriage of practical problem solving and psychological theorizing? *Applied Cognitive Psychology, 22*, 815–826.
- Earhart, B., La Rooy, D. J., Brubacher, S. P., & Lamb, M. F. (2014). An examination of “don’t know” responses in forensic interviews with children. *Behavioral Sciences & the Law, 32*, 746–761.
- Evans, J. R., & Fisher, R. P. (2011). Eyewitness memory: Balancing the accuracy, precision and quantity of information through metacognitive monitoring and control. *Applied Cognitive Psychology, 25*, 501–508.
- Eyewitness identification. (2017, January 16). In Wikipedia. Retrieved February 11, 2017 from https://en.wikipedia.org/w/index.php?title=Eyewitness_identification&oldid=760434277
- Fahsing, I. A., Ask, K., & Granhag, P. A. (2004). The man behind the mask: Accuracy and predictors of eyewitness offender descriptions. *Journal of Applied Psychology, 89*, 722–729.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.
- Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology, 74*, 722–727.
- Garrett, B. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Cambridge, MA: Harvard University Press.
- Granhag, P. A., Ask, K., Rebelius, A., Öhman, L., & Giolla, E. M. (2013). ‘I saw the man who killed Anna Lindh!’ An archival study of witnesses’ offender descriptions. *Psychology, Crime & Law, 19*, 921–931.
- Innocence Project. (2013). *Brief of amicus curiae the Innocence Project in support of Thomas and Raymond Highers’ motion in limine for an eyewitness identification expert*. Retrieved from https://www.innocenceproject.org/files/imported/highers_amicus_final.pdf
- Innocence Project. (2017). *Eyewitness misidentification*. Retrieved from <https://www.innocenceproject.org/causes/eyewitness-misidentification/>
- 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.
- Koehnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime & Law, 5*, 3–27.
- Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quality-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 36*, 297–315.
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Horowitz, D., & Abbott, C. B. (2007). Does the type of prompt affect the accuracy of information provided by alleged victims of abuse in forensic interviews? *Applied Cognitive Psychology, 21*, 1117–1130.
- Loftus, E., & Ketcham, K. (1994). *The myth of repressed memory: False memories and allegations of sexual abuse*. New York, NY: St. Martin’s Press.
- Loftus, E. F. (2003). Make-believe memories. *American Psychologist, 58*, 867–873.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31.
- Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*, 720–725.
- Mickes, L. (2015). Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in

- investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, 4, 93–102.
- Mosser, A., Fisher, R. P., Molinaro, P., Satin, G., & Manon, M. (2016, March). *Train-the-trainer: Training law enforcement investigators in the CI*. Paper presented at the Annual Conference of the American Psychology-Law Society, Atlanta, GA. Abstract retrieved from <http://ap-ls.wildapricot.org/resources/Documents/AP-LS2016/2016ConferenceProgram.pdf>
- National Research Council. (2009). *Strengthening forensic science in the United States: A path forward*. Washington, DC: National Academy Press.
- Newlin, C., Steele, L. C., Chamberlin, A., Anderson, J., Kenniston, J., Russell, A., & Vaughan-Eden, V. (2015). *OJDDP Juvenile Justice Bulletin. Child forensic interviewing: Best practices* (NCJ Publication no. 248749). Retrieved from <http://www.ojjdp.gov/pubs/248749.pdf>
- 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.
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged abuse victims. *Child Abuse & Neglect*, 24, 733–752.
- Paulo, R. M., Albuquerque, P. B., & Bull, R. (2016). The enhanced cognitive interview: expressions of uncertainty, motivation and its relation with report accuracy. *Psychology, Crime & Law*, 22, 366–381. doi:10.1080/1068316X.2015.1109089
- Penrod, S., & Cutler, B. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law*, 1, 817–845.
- Police and Criminal Evidence Act 1984* (2011). Code D. Retrieved from <https://www.gov.uk/government/publications/pace-code-d-2011>
- Rivard, J. R., Fisher, R. P., Robertson, B., & Hirn Mueller, D. (2014). Testing the cognitive interview with professional interviewers: Enhancing recall of specific details of recurring events. *Applied Cognitive Psychology*, 28, 917–925.
- Roberts, W. T., & Higham, P. A. (2002). Selecting accurate statements from the cognitive interview using confidence ratings. *Journal of Experimental Psychology: Applied*, 8, 33–43.
- Roediger, H. L., III, 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, NY: Oxford University Press.
- Snook, B., & Keating, K. (2011). A field study of adult witness interviewing practices in a Canadian police organization. *Legal and Criminological Psychology*, 16, 160–172.
- Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence–accuracy relation in eyewitness identification studies. *Psychological Bulletin*, 118, 315–327.
- Stebly, N. K., & Dysart, J. E. (2016). Repeated eyewitness identification procedures with the same suspect. *Journal of Applied Research in Memory and Cognition*, 5, 284–289.
- Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* (NCJ Publication no. 178240). Washington, DC: United States Department of Justice, Office of Justice programs. Retrieved from <https://www.ojp.usdojncjrs.gov/pdffiles1/nij/178240.pdf>
- Thompson, M. B., Tangen, J. M., & McCarthy, D. J. (2014). Human matching performance of genuine crime scene latent fingerprints. *Law and Human Behavior*, 38, 84–93.
- Thompson, W. C. (2013). Forensic DNA evidence: The myth of infallibility. In S. Krinsky & J. Gruber (Eds.), *Genetic explanations: Sense and nonsense* (pp. 227–255). Cambridge, MA: Harvard University Press.
- Thompson, W. C., Ford, S., Doom, T., Raymer, M., & Krane, D. E. (2003). Evaluating forensic DNA evidence: Essential elements of a competent defense review. *The Champion*, 27, 16–25.
- U.S. Federal Bureau of Investigation, Department of Justice. (2011). *The FBI quality assurance standards audit for forensic DNA testing laboratories*. Retrieved from <https://www.fbi.gov/file-repository/qas-audit-for-forensic-dna-testing-laboratories.pdf/view>
- van Koppen, P. J., & Lochun, S. K. (1997). Portraying perpetrators: The validity of offender descriptions by witnesses. *Law and Human Behavior*, 21, 661–685.
- Wells, G. L., Small, M., Penrod, S., Malpass, R., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photo-spreads. *Law and Human Behavior*, 22, 603–647.
- Wells, W., Campbell, B., Li, Y., & Swindle, S. (2016). The characteristics and results of eyewitness identification procedures conducted during robbery investigations in Houston, TX. *Policing: An International Journal of Police Strategies & Management*, 39, 601–619.
- Wixted, J. T., Mickes, L., Clark, S., Gronlund, S., & Roediger, H. (2015). Confidence judgments are useful in eyewitness identifications: A new perspective. *American Psychologist*, 70, 515–526.
- 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, USA*, 113, 304–309.
- 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.
- Woolnough, P. S., & MacLeod, M. D. (2001). Watching the birdie watching you: Eyewitness memory for actions using CCTV recordings of actual crimes. *Applied Cognitive Psychology*, 15, 394–411.
- Yates, S. Q. (2017, January 6). *Memorandum for heads of department law enforcement components, all department prosecutors*. Washington, DC: Office of the Deputy Attorney General, U.S. Department of Justice. Retrieved from <https://www.justice.gov/archives/opa/press-release/file/923201/download>
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71, 291–301.