

Chapter 6: Eyewitness Evidence

Melissa F Colloff and Heather D Flowe

School of Psychology, University of Birmingham

To appear in

Davies, Beech, & Colloff (Eds).

Forensic Psychology:

Crime, Justice, Law, Interventions (4th Edition)

Wiley-Blackwell

Learning outcomes

By the end of this chapter, you should be able to:

- Understand why a psychological analysis of eyewitness evidence is important in criminal investigation and the legal process
- Appreciate the principal research methods used to explore eyewitness memory and evidence
- Understand the implications of existing research findings for the conduct of investigators and the courts

Introduction

Imagine that you were an eyewitness to a criminal event. Perhaps you saw a suspicious man in an area where you later learn a child has been abducted, or you caught a glimpse of the young man when he grabbed the bag from your arm. The investigating police officers may ask you to describe what you saw and the appearance of the culprit. In some jurisdictions, such as the UK, you may also have a formal video-recorded interview with the police. Investigations often hinge on the quality of statements given by a witness or victim, because other evidence—such as DNA, physical injury, or CCTV—may be lacking, even for interpersonal types of crimes such as robbery or sexual assault (e.g., Kelly et al., 2005). You may also be asked to attempt to identify the culprit from a lineup. In this lineup, the police officers place their suspect (who is either innocent or guilty) among other known-to-be-innocent lineup members, called fillers. If you make a positive identification of the suspect—you say: “That’s him!”—this is likely to be interpreted as compelling evidence of guilt. Suspects who have been identified are more likely to be found guilty at court than those who have not (e.g., Devlin, 1976). In short, eyewitness evidence plays a critical role in how cases proceed through the Criminal Justice System.

The influence of eyewitness evidence is, however, perhaps concerning when we consider that memory reports can be inaccurate (see Case Study 1). When poor memory collection procedures are used, eyewitnesses can include incorrect information that has been introduced

from others (e.g., Gabbert et al., 2003), become more certain of their accounts when they are given positive feedback (Stebay et al., 2014), and even develop false memories for entire events that did not happen (e.g., Loftus & Pickrell, 1995). Eyewitnesses can also make erroneous identifications, and weak identification evidence can be given too much weight in legal decision-making. In the United States since 1989, 365 convictions have been overturned based on new DNA evidence. The Innocence Project estimates that over 70% of these wrongful convictions involved misleading eyewitness evidence, and in some of these cases an eyewitness misidentified an innocent suspect in a lineup (Innocence Project, 2019)¹. Another type of identification error, failing to identify the culprit when he is in the lineup, can result in the real culprit being free to commit additional crimes.

Because collecting accurate evidence is important, psychological scientists have long studied eyewitness memory in laboratory conditions. Researchers usually use a mock-crime methodology. Subjects watch a staged crime—sometimes live, but usually one that has been videotaped—then, after a delay, are asked to recall what happened (i.e., provide testimony) or if they can identify the culprit from a lineup. In a lineup study, some subjects are presented with a lineup in which the culprit is present (target-present lineup), and the remaining subjects are presented with a lineup in which the culprit is absent (target-absent lineup). Target-present lineups represent the real-world situation in which the police suspect is guilty, whereas target-absent lineups represent the situation in which the police suspect is innocent. In both target-present and target-absent lineups, subjects can make one of three possible identification responses: they can identify the suspect, a filler, or state that the real culprit is not present. In real life criminal investigations, only incorrect identifications of innocent suspects result in erroneous criminal proceedings being brought against that person, because identifications of fillers are known errors.

It is important to consider how actual crimes differ from the events that psychologists use in the laboratory to make inferences about real eyewitnesses. The unique nature of witnessing a real crime might influence how well events are remembered. In a study of a fatal shooting in Canada, Yuille and Cutshall (1986) investigated the accuracy of witness' statements recorded

¹ Note that many of these cases involved eyewitness factors that were not due to memory error, such as the witness stating that they saw the suspect at the scene but did not actually report seeing them commit the crime, or the witness erroneously naming a suspect they knew was innocent.

during initial police interviews at the time of the crime and again 4 to 5 months later. Not only were the witnesses' recollections overall very accurate, but accuracy did not decline over time across interviews. These findings might seem surprising considering the studies described in subsequent sections, so it is important to bear in mind the possibility that real-life crimes could be better remembered than laboratory studies would have us believe (Koriat & Goldsmith, 1994). It is certainly important to exercise caution before automatically assuming that findings generalise from the lab to the real world (Wagstaff et al., 2003).

Case Study 1: The Andrew Malkinson Case

In July 2003 in Greater Manchester, a 33-year old woman was walking home in the early hours of the morning after having an argument with her partner. She was approached by a man at the roadside and beaten and raped. Two weeks after the attack, the woman identified Andrew Malkinson from a video identification parade. At court, the victim stated that she was more than 100% certain that Malkinson was her attacker. He was charged with rape and spent 17 years in prison. But, to this day, Malkinson maintains his innocence. It was not reported how certain the victim was when her memory was tested at the identification parade, so it is not clear if her confidence was inflated by the time she testified in court. Moreover, there were no reasonable grounds for suspicion before Malkinson was placed in the parade: Malkinson became a suspect because officers that attended to the victim after the attack thought her description sounded like Malkinson, who they remembered from a routine car check a few days earlier. In 2022, new DNA testing techniques showed that Malkinson's DNA was not at the scene; but, DNA from another male was present. At the time of this writing, Malkinson's conviction is being reconsidered by the Criminal Cases Review Commission. The case highlights some potential issues in the way that legal decision-makers decide to collect and interpret eyewitness evidence.

The Process of Remembering

Imagine you are being interviewed in a police station about the crime you witnessed. You are in an unfamiliar place, feeling under pressure to recount the event as accurately and in as much detail as possible. The police ask you specific questions about details, some of which you might struggle to remember. When explaining how your memory might work in this situation,

researchers often refer to three stages: *encoding*, *storage*, and *retrieval*. At each stage, a number of factors can operate to undermine or support the quality and accuracy of memory.

The first stage of memory is *encoding*, and involves the creation of a memory. The quality of encoding can be affected by the extent to which the witness had paid attention to the crime, or how near to the crime they were, for example. Other encoding-based factors are reviewed in the estimator variables section of this chapter.

The second stage is *storage*, or memory retention. Memory is stored for later retrieval after it is encoded. Even if the witness encoded the memory under pristine conditions (i.e., conditions that lead to a strong memory), their memory accuracy might be influenced by hearing other people describe the attack. A large body of research into false memory formation, for example, shows that stored memories can be vulnerable to external influences, and the resulting memory representations can include a mixture of observed information as well as information acquired after the crime from other sources (e.g., media accounts, other witnesses). This phenomenon, termed the misinformation effect, is explained in more detail later in the chapter.

Retrieval describes the process by which memory is accessed, and in legal contexts, involves the witness explaining what happened during the event, as well as their taking an identification test, such as a lineup. Strong memories often come to mind easily, but other memories might be difficult to access. Closing your eyes while relaying information may increase the accuracy of memory reports (see Vredeveltdt et al., 2011). Other ways in which the accuracy of retrieval can be affected are covered in Chapter X.

Factors influencing witness memory accuracy

Much research has focused on understanding the factors that might make memory evidence more or less accurate. There are two types of factor affecting eyewitness accuracy: *estimator variables* and *system variables* (Wells, 1978). *Estimator variables* cannot be controlled by the legal system, so their impact has to be estimated by legal professionals. Estimator variables are primarily related to the encoding and storage stages of memory, particularly the characteristics of the crime event. *System variables* can be controlled by the legal system and are primarily related to the retrieval stage of memory. Table 1 displays example estimator and system variables. Understanding the influence of estimator variables is important to evaluate the likely performance

of eyewitnesses under various conditions, while understanding system variables enables recommendations about best practice to be made.

Table 1. Example System and Estimator Variables that can affect Eyewitness Testimony and Identification

	Testimony	Identification
System Variables	• Questioning format (free-recall, closed, leading) • Interview approach (e.g., Cognitive Interview) • Interviewer rapport with the witness • Exposure to post-event misinformation • Planning the structure and content of the interview • Complex and compound questions or language • Using an interview approach appropriate for the witness's cognitive ability • Co-witness discussions	• The lineup administrator knows who is the suspect • The appearance of the lineup fillers • Showups or lineups • Presentation modality (e.g., live, photos, videos) • Repeated identification attempts with the same suspect • Sequential or simultaneous presentation • Immediate collection of witness confidence at the initial memory test • Lineup instructions
Estimator Variables	• Intoxication • Delay between the crime and memory recall or lineup • Distance from the culprit at encoding • Witness and culprit of different or same ethnicities • Delay between the crime and the identification procedure • Witness age • Exposure time to the culprit at encoding • Witness stress	

- The presence of a weapon
- Witness prior familiarity with the culprit
- Illumination

This section reviews decades of research highlighting that there are many estimator and system variables that can impair memory accuracy. Three aspects of memory performance are considered: memory accuracy (i.e., the ability of the witness to accurately distinguish between correct and incorrect information they report about the crime), memory reliability (i.e., the probability that information recalled by the witness at a given level of certainty is correct), and completeness (i.e., the quantity of information reported by the witness).

From the outset, it is important to note that legal systems globally are often concerned with spotting the presence of estimator variables to make a distinction between good and poor eyewitness evidence. In England, for example, following a landmark ruling in the Appeal Court (*R v Turnbull*, 1976), when identity is disputed a trial judge must advise the jury to consider the circumstances of the identification (see Box 1).

Box 1. The Turnbull Guidelines (*R v Turnbull and Others*, 1976)

In a case that relies substantially on disputed eyewitness identification evidence, the trial judge must warn the jury about the need for caution before relying on identification evidence to convict the defendant. The judge should instruct the jury that it is possible that a convincing witness, or several witnesses, may be mistaken, and to carefully consider the circumstances of each witness's identification. The relevant circumstances are summarised by the acronym ADVOKATE:

1. Amount of time the perpetrator was in view
2. Distance of the witness from the perpetrator
3. Visibility of the perpetrator (e.g., lighting)
4. Obstruction to the witness's view?
5. Known to the witness?
6. Any reason the witness would remember the suspect?
7. Time delay between the incident and the identification procedure

8. Error. Are there discrepancies between the description given to the police at the time of the incident and the appearance of the suspect?

One conclusion that might be drawn from the legal guidelines and the research that underpins them might be that eyewitnesses should not be relied upon when the encoding conditions were poor. Yet, more recent research suggests that eyewitness memory may be more reliable than previously believed, such as when witness certainty is considered (see Case Study 2). When researchers collect witness confidence (e.g., “How sure are you that you are correct?”), identifications and memory reports made with high confidence are typically highly accurate (e.g., Wixted & Wells, 2017; Wixted et al., 2018). Participant witnesses are less likely to make a positive identification with high confidence when the crime was witnessed under poor conditions. But, when they do, the identification is highly likely to be accurate. Similarly, when participants freely report their memories to an interviewer, the accuracy of the information they report tends to be very high. This is because witnesses tend to report information when they are sure it is accurate and withhold information when they are unsure it is accurate. Therefore, the quantity of information reported may be reduced when the encoding conditions are relatively poor, but the accuracy of the reported information remains high. In the next sections, we consider the effects of example estimator and system variables on (1) memory reporting accuracy and completeness with regard to witnesses’ statements and testimony, and (2) identification accuracy and confidence.

Witness Statements and Testimony

Estimator Variables

Intoxication.

A victim, suspect and/or witness may have been under the influence of alcohol or other psychoactive drugs around the time of the crime (Evans et al., 2009). Alcohol intoxication is particularly common in sexual and physical offences because these crimes are more likely to occur at night when substance use is more common than other times of day (Office of National Statistics, 2011/2012). Alcohol affects regions of the brain (e.g., the hippocampus) that consolidate newly

learned information into long-term memory (White, 2003). Two different types of blackout can result: an en bloc blackout or a fragmentary blackout (Goodwin et al., 1969a). With an en bloc blackout, a person has no memory and cannot recall anything about the crime for whole stretches of time. Memory loss is less in a fragmentary blackout and can be successfully recalled in part (Goodwin et al., 1969b). In addition to alcohol's physiological effects on memory consolidation, people's preconceptions about the effect alcohol has on memory can be more detrimental than the alcohol itself (Assefi & Garry, 2003). However, this finding has not been consistently reported. Clifasefi et al. (2006) found that whilst all participants were told that they were drinking alcohol, only participants who actually drank alcohol failed to notice an unexpected event, while those who were sober, did notice it.

Accuracy or Completeness? Research highlights differences between alcohol intoxicated and sober participants in the completeness and accuracy of the memories they report. Flowe et al. (2016) found that although participants who were intoxicated during the encoding of a sexual assault scenario provided significantly less information overall than their sober counterparts, the accuracy of the information they provided did not differ one day or four months after the crime. This suggests that people who witness crimes whilst intoxicated are less likely than their sober counterparts to answer questions when they are unsure about the accuracy of the information they remember. In a follow up study, participants were interviewed about the sexual assault using police interview procedures one week later, and women who were alcohol intoxicated versus sober during encoding provided less information, but it was just as accurate (Flowe et al., 2019). This research has important implications for understanding the accuracy of victims' accounts of actual sexual assaults. Moreover, a meta-analysis of the literature on alcohol intoxication during encoding and witness accuracy found that intoxication at encoding decreases the amount of information reported, but does not decrease accuracy (Jores et al., 2019).

Delay.

The delay, or retention interval, between the crime and when witnesses provide statements and testimony can vary from a few minutes to several years. All other things being equal, information is more likely to be forgotten with increasing delay. For example, Van Koppen and Lochum (1997) performed an archival analysis of eyewitness descriptions of commercial bank robbers and found that accuracy and completeness decreased the longer the delay between the

crime and when the description was given by the eyewitness. Memory reports may be especially accurate when witnesses are interviewed immediately after the crime, because memory traces are the strongest, most accessible, and most available (Tuckey & Brewer, 2003). Additionally, both in the cognitive memory literature (e.g., Wixted & Ebbesen, 1991), and the eyewitness memory literature (e.g., Tuckey & Brewer, 2003), memory initially decays rapidly, followed by a much slower decline. However, real-world eyewitnesses might recall the critical incident on several occasions before they provide testimony. This can affect the rate of forgetting over time. The amount and accuracy of information forgotten can be minimized by questioning witnesses relatively soon after the crime (e.g., Hope et al., 2014).

Accuracy or Completeness? Witnesses tend to decrease the amount of information they report with increasing delay, and this helps to maintain the overall accuracy of the information they report over time. As an example, Ebbesen and Rienick (1998) had participants interact briefly with a stranger. After 28 days they tested the participants' memories for what the other person looked like. The number of correct descriptors that participants recalled did not decrease over the 28 days among participants who had their memory repeatedly tested several times during the 28-day period. In contrast, if participants' memory was tested for the first time after a delay of 7 to 28 days, the number of correct details decreased as the length of time between encoding the stranger and initial testing increased. The number of incorrect details reported did not increase over time, regardless of when memory was initially tested.

Witnesses can also vary the completeness of their statements by regulating the specificity, or coarseness, of the information they report. The regulation of specificity helps to maintain accuracy. For instance, with increasing retention interval length, witnesses may be more likely to report that the getaway car was dark in colour rather than navy blue (see Goldsmith et al., 2005). Accuracy is maintained as specificity decreases, because information is more likely to be accurate if it is reported at a coarser level.

System Variables

Question Format.

The procedure used to gather information from witnesses is a key variable that can affect the accuracy of statements and testimony. For example, the format of the questions that

eyewitnesses are asked is important. Questions that have an open-response format (e.g., “Describe the culprit”) allow the participant to freely report what they remember. In contrast, witnesses’ responses are limited by closed-format questions (“What colour was the culprit’s shirt?”) as well as yes/no questions (“Was the culprit’s shirt red?”). The accuracy and completeness of answers can vary depending upon question format (e.g., Davies et al., 2000; Fisher et al., 1987). Leading questions suggest to witnesses that an answer to a question is required and can affect the accuracy of their response (e.g., Loftus & Palmer, 1974).

Completeness or Accuracy? Witnesses are able to determine what information to volunteer during a police interview when they are allowed to freely recall their memories. Research has found that explicitly allowing witnesses the option to answer “I don’t know” during questioning increases the accuracy of the information reported (Weber & Perfect, 2013). Theoretically, having the explicit option to withhold reporting information prompts people to monitor and control their memory reports, leading them to report information when they are more certain of its accuracy. Laboratory research has found that when participants volunteer less information overall about a mock crime they witnessed, the information they report is more accurate than if they had tried to provide answers to all of the questions that they were asked. This pattern of responding is known as the informativeness accuracy trade-off (Weber & Brewer, 2008; also see Koriatic & Goldsmith, 1996).

Memory Retrieval Tools.

Resource limitations, such as staff shortages, can make it difficult for the police to interview all of the witnesses who saw a crime soon afterwards. As discussed previously, memory is more complete when the retention interval between the crime and the interview is relatively short. An innovation developed to help the police interview as many witnesses as possible soon after the crime is the Self-Administered Interview (SAI) (Gabbert et al., 2009). The witness uses the SAI to perform a self-guided interview. The SAI entails interview prompts presented in a written booklet. Witnesses perform the SAI as soon as possible after the crime and write their responses to the prompts in the booklet. The tool was also developed to improve the accuracy of memory retrieval. The SAI includes mnemonic devices to help witnesses better remember, such as instructing witnesses to reinstate, or mentally rehearse, the context in which the crime occurred. The SAI is based on the principles that underpin the Cognitive Interview (Fisher & Geiselman,

1992), which is discussed in more detail in Chapter X. Another example of an innovative tool that can help witnesses better remember information is the timeline technique (Hope et al., 2013). It was developed to aid witnesses by establishing the temporal context of the event, which can improve accuracy.

Completeness or Accuracy? An important consideration in developing tools to improve witness memory retrieval, is the extent to which the tool affects completeness as well as accuracy. Procedures, such as the Cognitive Interview, that instruct witnesses to report everything in as much detail as they can remember may lead witnesses to report more details; but, this may also come at the expense of them reporting more incorrect details than otherwise would be the case (Memon et al., 2010). This is because such instructions decrease witnesses' memory reporting standard, or how sure they must be about the likely accuracy of the information before they report it to the interviewer (see Memon & Higham, 1999). Information is less likely to be accurate as witnesses lower their reporting standard.

Minimizing Exposure to Misinformation.

Witnesses to the same crime can pay attention to, remember, and misremember, different aspects of it. Through subsequent discussion with each other, witnesses might learn information about the event that they did not actually observe. If witnesses incorporate this information into their own testimony, this can be problematic for memory. The *misinformation effect* refers to when the accuracy of a witness' memory report is decreased by their exposure to incorrect information about the crime (Loftus & Hoffman, 1989). While exposure to misinformation can also be considered an estimator variable, it is possible for law enforcement to minimize its occurrence by interviewing witnesses alone, and by instructing witnesses not to discuss the case with others or read media reports about the case.

In their seminal paper on the misinformation effect, Loftus et al. (1978) presented participants with a slide sequence that depicted a simulated accident involving a car and a pedestrian. Afterwards, participants answered 20 questions concerning the event. For experimental participants, one of the questions suggested misleading information about the type of traffic sign at the scene of the accident (it described it as a Stop sign when in fact it was a Yield sign or *vice versa*). Control participants received consistent information about the type of traffic sign on the questionnaire rather than misleading information. A two-alternative forced-choice recognition test

was given in the final phase of the experiment, featuring two slides, one showing the correct and the other showing the misleading sign. Compared to controls, experimental participants were far more likely to choose the incorrect slide, suggesting that their memory of the accident had been altered by exposure to the misleading post-event information.

How do we come to conform our memory reports to another person's? There are a few theoretical explanations. According to the *updateable memory hypothesis*, the misinformation effect occurs because stored memories are overwritten with the misinformation we encounter (e.g., Loftus, 1979; Wright & Loftus, 1998; Braun & Loftus, 1998). According to this account, memory is the product of reconstruction, and we continually reinterpret events that we see. To illustrate, a witness in the Loftus et al. (1978) experiment who accepts the misinformation embedded in the questionnaire might visualize the misleading detail as part of the scene. The witness will then assimilate the fragments of misinformation into their memory of the event. When they come to recall it, their memory will be based on all of the fragments, misinformation included. The *strategic effects* account proposes that memory is not impaired by misinformation, but rather participants' memory reports are affected by task demands and test-taking strategies (McCloskey & Zaragoza, 1985a, 1985b; Zaragoza et al., 1987). The *blocked memory access* account proposes that memory traces for the original event and the misleading information coexist in memory. There is response competition between the traces, and whether the originally witnessed information is retrieved from memory depends on whether memory for the misinformation is relatively stronger and blocks memory retrieval (e.g., Bekerian & Bowers, 1983; Chandler, 1991). The *source monitoring account* proposes that people sometimes have difficulty distinguishing the source of their memory and will mistakenly report misinformation as something they actually had witnessed (Lindsay & Johnson, 1987).

Completeness or Accuracy? Warning witnesses about misinformation before the interview can influence them to adopt a stricter standard for reporting information to the interviewer. Further, helping witnesses consider the source of their memories can enable them to withhold misinformation on memory tests. For example, Blank and Launay (2014) performed a meta-analysis of misinformation effect studies and found that warning people about the deleterious effects of misinformation on memory reporting boosts accuracy by as much as 50% relative to control conditions. Although one might be concerned that such warnings might also decrease the

number of correct details that witnesses report to interviewers, they argued that such costs were outweighed by the benefits of ameliorating misinformation reporting.

Case Study 2: Edinburgh Fringe Assault

In August 2017, a 19-year-old teenager had been at Edinburgh Fringe festival and left a nightclub at around 5am, deciding to walk home. An older man helped her into his campervan which was parked in the street. The teenager described: "He bent down and picked me up and took me towards his vehicle. I was quite drunk. I was struggling to get into the van. He got in behind me. There was a mattress there and I almost immediately fell asleep." Sometime later, the teenager woke up to find the man raping her. He then drove her a short distance, where she ran away. A taxi driver who saw the teenager running said that she was very distressed. The teenager's testimony resulted in the apprehension of Mark Adams—a civil servant who happened to be private secretary to both John Major and Tony Blair during the 1990s. The case was heard in 2019, and, after spending two hours in deliberations, jurors returned a guilty verdict to one charge of rape. Adams was subsequently convicted for another rape in 2015, and in 2018 for sexual assault. This case illustrates how a victim, despite being distressed and alcohol intoxicated, was able to provide accurate testimony that enabled the detention of a serial sex offender.

Eyewitness Identifications

Estimator Variables

First, it is important to note that confidence was previously believed to not be informative of accuracy on eyewitness identification tasks. This is because the approach often used to analyse the data underestimated the relationship between confidence and accuracy (see Box 2). Here, we provide an overview of some of the latest laboratory research examining estimator variables with respect to identification accuracy and reliability, which is the probability that an identification made at a given level of confidence is correct.

Box 2: Measuring The Relationship Between Eyewitness Confidence and Accuracy

To measure the relationship between eyewitness confidence and accuracy, much prior research calculated the correlation coefficient, which can range from 0 (no correlation) to 1 (perfect correlation). Yet, a small correlation coefficient does not necessarily indicate a poor relationship between confidence and accuracy (Juslin et al., 1996). To compute a correlation coefficient, imagine a graph where confidence (1, 2, 3...) is plotted on the x-axis and accuracy (where 0 = incorrect and 1 = correct) on the y-axis (see Fig 1a). Each point on the plot represents one witness (note that there are many points that cannot be seen because they fall directly on top of each other). Correlation coefficients fit a straight line through these data. The distribution of confidence judgments across the confidence scale heavily influences the line. In experiments, confidence judgments often fall within a restricted range and computing a correlation coefficient in this case underestimates the relationship between confidence and accuracy (Lindsay et al., 1998).

A more suitable technique for testing whether people can assess the accuracy of their memories is to plot *average* accuracy at different levels of confidence—that is, plot confidence-accuracy curves (see Fig 1b). This technique is unaffected by the distribution of confidence judgments because average accuracy at a particular level of confidence is the same regardless of how many identifications are made with that level of confidence (Brewer et al., 2002; Mickes, 2015).

Why does the analysis approach matter? Using the correlation coefficient can lead to misleading conclusions. In a recent review, Wixted and Wells (2017) calculated the mean correlation coefficient to be .37 across three seminal studies. As a result of the early research on confidence and accuracy using the correlation coefficient, legal systems globally came to disregard eyewitness confidence as an indicator of accuracy. This is unfortunate because when Wixted and Wells (2017) used confidence-accuracy curves to re-analyse the data they found that, on average, accuracy was 61% correct at low confidence and 97% correct at high

confidence. They also reviewed 20 additional studies that tested fair lineups and the strong relationship between confidence and accuracy is evident in each of the studies.

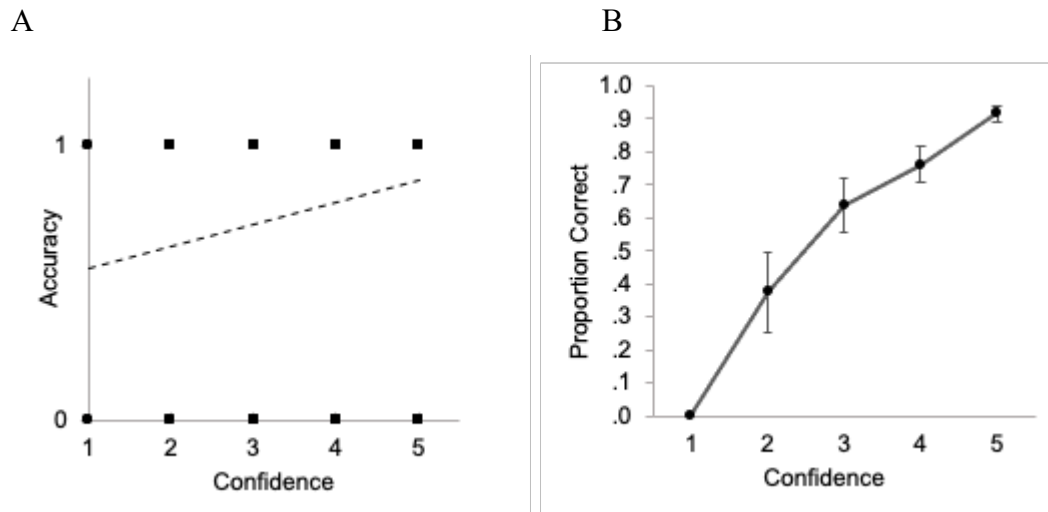


Figure 1. Examining the relationship between confidence and accuracy by (a) computing the correlation coefficient by fitting a straight line through confidence and accuracy data, and (b) plotting confidence-accuracy curves.

Distance.

People seem to be able to accurately recognise familiar faces, like celebrities or family members, from substantial distances (e.g., Greene & Fraser, 2002). Eyewitnesses, however, are often required to identify perpetrators that they have only seen once before. Eyewitnesses tend to be less accurate in recognising a perpetrator if they viewed him committing the crime from farther away (e.g., Lindsay et al., 2008; Nyman, Antfolk et al., 2019). Increased distance decreases correct identifications, increases false identifications, and reduces people's ability to discriminate guilty from innocent suspects (Lockamy et al., 2020; Semmler et al., 2018). One explanation is that details about the features of faces become increasingly coarser as the witness moves further away from the perpetrator (Loftus & Harley, 2005).

Confidence? Participants are typically highly accurate when they are very confident, regardless of viewing distance. Identifications made with high confidence have been shown to be similarly accurate at short and long distances (Semmler et al., 2018), even up to 40 meters (Nyman, Lampinen et al., 2019). Lockamy et al. however, found that accuracy was lower at

longer viewing distances (20 meters versus 10 or 3 meters), even at high confidence. Participants at 20 meters in this study performed poorly on the identification task which may have limited the extent to which confidence predicted accuracy. Overall, it appears that a confident witness could be relied upon to make an accurate suspect identification even if they were far away from the perpetrator during the crime. However, even confident identifications made at very long distances, or when memory is otherwise very poor, might not be as reliable.

Ethnicity.

Much research illustrates that people are less able to accurately perceive and remember faces that are of a different race or group than their own; a phenomenon called the *Own Race Bias* (ORB; Malpass & Kravitz, 1969). A meta-analysis found that identifications of other-race compared with own-race faces are 56% more likely to be erroneous (Meissner & Brigham, 2001). The ORB is thought to occur due to cognitive and social processes at encoding. For example, the *dual-process* account proposes that observers are more likely to encode more information, create a more diagnostic representation, and retrieve specific episodic information that aids accurate recognition for same-race than other-race faces (Meissner et al., 2005). The *in-group/out-group* model proposes that, whereas people attend to and encode configural properties of own-race faces, other-race faces are automatically categorized and processed on a relatively shallow level (e.g., Sporer, 2001). Multiple other socio-cognitive mechanisms may contribute to the ORB.

Confidence? Confidence appears to be informative about identification accuracy, for both own-race and other-race identifications. In some experiments, participants are tested on their memory for multiple faces, with multiple lineups. Generally speaking, in these studies, identifications made with high confidence have been shown to be similarly accurate for own- and other-race faces (Colloff et al., 2022 Experiment 1; Dodson & Dobolyi, 2016; Nguyen et al., 2017), though other-race identifications were slightly more overconfident in Dodson and Dobolyi's (2016) study. In real life, however, witnesses typically view one crime and take one lineup. In a second experiment by Colloff et al. (2022), participants only saw one perpetrator in a mock-crime and lineup. Here, high-confidence other-race identifications were less likely to be accurate than high-confidence own-race identifications. Together, these findings suggest that in multiple-trial experiments the race of the witness and culprit appears to be minimally important when examining the confidence-accuracy relationship, but the ORB may be an important factor for legal decision-

makers when the witness is tested only once. More research is required. Nevertheless, and notably, for both own- and other-race identifications across studies, increased confidence is associated with increased suspect identification accuracy, highlighting that confidence can be a useful (but imperfect) indicator of likely accuracy.

Delay.

Face recognition ability typically declines over time for people who are not familiar to us. Increased time between encoding and the lineup test can decrease correct identifications, increase filler identifications, increase false identifications of innocent suspects, and reduce people's ability to discriminate guilty from innocent suspects (e.g., Juslin et al., 1996; Palmer et al., 2013; Sauer et al., 2010). Most studies test memory immediately and again around 1-3 weeks later, but one study tested memory up to 9 months after encoding (Read et al., 1998). There are several explanations of why forgetting happens over time (e.g., Wixted, 2022). There may be random error in the structural changes to the synapse that occur after learning, known as decay. Other information may compete or degrade the to-be-remembered information, known as interference. Or, forgetting might occur because the context at the time of retrieval differed from the context at the time of encoding.

Confidence? Participants typically seem to be highly accurate when they are highly confident, regardless of the delay between the event and lineup (Wixted, Read, et al., 2016). Identifications made with high confidence have been shown to be similarly accurate for lineup identification decisions made immediately compared to after 1 week (Palmer et al., 2013), 1 hour compared to 1 week (Juslin et al., 1996); immediately compared to around 3 weeks (Sauer et al., 2010), and even 3 months compared to 9 months (Wixted, Read et al., 2016). Overall, this research suggests that a confident witness could be relied upon to make an accurate suspect identification even after a delay.

Age.

Older people make fewer correct responses in tests of face recognition (e.g. Bartlett & Fulton, 1991). On lineup tasks, older witnesses aged over 60 years find it harder to discriminate between innocent and guilty lineup members compared to younger (aged 18-30) and middle-aged witnesses (aged 31-59; Colloff et al., 2017). Older adults can also make both fewer correct

identifications and more mistaken identifications (e.g., Searcy et al., 1999, 2000). Some researchers have suggested that the age-related decline in accuracy occurs because older adults are more willing to choose a person in the lineup, resulting in an increased rate of false identifications (e.g., Sporer & Martschuk, 2014; Wilcock et al., 2005). Others have suggested that changes in memory function can explain poorer performance with age (e.g., Colloff et al., 2017). For example, according to dual-process accounts of memory, recognition is based on two processes: recollection and familiarity (see Yonelinas, 2002, for a review). Recollection involves retrieving specific contextual information about the original stimulus, such as source, time, place, and thoughts, whereas familiarity is a sense that the stimulus has previously been encountered without contextual detail. Older adults seem to have deficits in recollecting diagnostic contextual information and therefore rely more on less diagnostic familiarity processes than their younger counterparts and this tendency promotes memory errors (e.g., Searcy et al., 1999).

Research also suggests that face recognition abilities improve throughout childhood (Chung & Thomson, 1995; but see also Crookes & McKone, 2009). On lineup tasks, younger children find it harder to discriminate guilty from innocent lineup members than older children and young adults (e.g., Humphries & Flowe, 2015; Fitzgerald & Price, 2015). Again, there is debate regarding the underlying mechanisms. Some researchers suggest that younger children are more willing to (correctly or incorrectly) choose a person in the lineup because they have difficulty withholding an inappropriate response (e.g., Dunlevy & Cherryman, 2013; Pozzulo & Lindsay, 1998). Other researchers instead suggest that improvements in identification performance with age are due to memory mechanisms gradually maturing (e.g., Brewer & Day, 2005; Winsor et al., 2021).

Confidence? Confidence can sometimes be informative about likely identification accuracy, for older adults and children, but the field is not yet in agreement. Older adults frequently make high-confidence errors (e.g., Dodson & Krueger, 2006). But other studies find that older adults tend to assign lower confidence ratings to their identification decisions on average than young adults, possibly suggesting that older adults are aware that they are less accurate (e.g., Memon et al., 2002; Neuschatz et al., 2005). One large-scale study found promising results for older adults (Colloff et al., 2017). Confidence increased as identification accuracy increased, regardless of age. The authors concluded that high-confidence identifications made by older adults

may be likely to be accurate, but may not be quite as accurate as high-confidence identifications made by younger adults.

Considering children, the previous consensus was confidence judgments do not reflect memory accuracy in children younger than 12 on lineup tasks, because children can be overconfident (e.g., Keast et al., 2007). However, the broader developmental literature has tested children using different tasks and paints a more positive picture of children's abilities (e.g., Sodian et al., 2012). These tasks usually ask children to view a sequence of stimuli (e.g., objects, faces) and later present children with some previously seen stimuli and some new stimuli and ask if they have seen each before. A recent reanalysis of the developmental and lineup literatures concluded that confidence provided considerable information about memory accuracy, from at least age 8, but possibly younger, because accuracy increased with confidence (Winsor et al., 2021). Children over the age of 10 were highly accurate when they were highly confident. The authors concluded that high-confidence identifications made by children from age 8 may be likely to be accurate, but they may not be as accurate as high-confidence identifications made by older children or adults.

Field and Archival Studies of Estimator Variables

Most of the research described above was conducted under controlled experimental conditions using mock crimes. A key question is whether estimator variables also influence the outcome of identifications made by real eyewitnesses. Archival studies involve examining records of previous cases, and field studies involve collecting data on eyewitness performance for current cases. The advantage of mock-crime studies is that the experimenter knows if a suspect identification is correct. In archival and field studies it is not known if the suspect identification is correct.

In one UK archival study, Valentine et al. (2003) examined the effect of estimator variables on the outcome of identifications made by approximately 600 witnesses who viewed over 300 live lineups. The suspect was more likely to be identified if the witness: was younger than 30; gave a detailed description; viewed the culprit at the scene for over a minute; and made a fast decision at the lineup. Lineups conducted within seven days of the incident produced a higher rate of suspect identifications (65%) than lineups conducted after eight days or more (38%). The ethnicity of the suspect and the witness (i.e., ORB) and whether a weapon was used during the incident was not associated with the rate of suspect identifications.

In another archival study, conducted in the US, Behrman and Davey (2001) examined the outcome of a non-random sample of 289 photographic lineups. Consistent with the UK data, lineups conducted within seven days of the incident produced a higher rate of suspect identifications (64%) than lineups held after eight days or more (33%), and the presence of a weapon was not associated with rate of suspect identifications. In contrast to the UK data, there was an ORB effect, whereby witnesses were more likely to identify the suspect when they were of the same ethnicity.

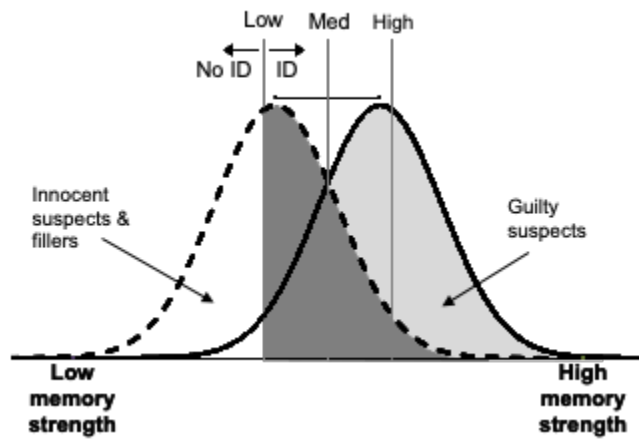
Does confidence predict accuracy for real eyewitnesses? One US field study of investigations in the Robbery Division of the Houston Police Department recorded eyewitness decisions made on 348 six-person photo lineups (Wixted, Mickes, et al., 2016). Eyewitnesses who made a positive identification were asked to provide a confidence rating on a 3-point scale. Most suspect identifications were made with high confidence, whereas most filler identifications were made with low confidence. High-confidence identifications appear to have been highly accurate. Using an advanced estimating technique, the authors estimated that high-confidence suspect identifications were 96% correct and low-confidence suspect identifications were around 50% correct. This field study provides compelling evidence that highly confident suspect identifications are likely to be accurate, despite the presence of real-world estimator variables that may influence memory encoding and storage.

A Theoretical Explanation of Confidence

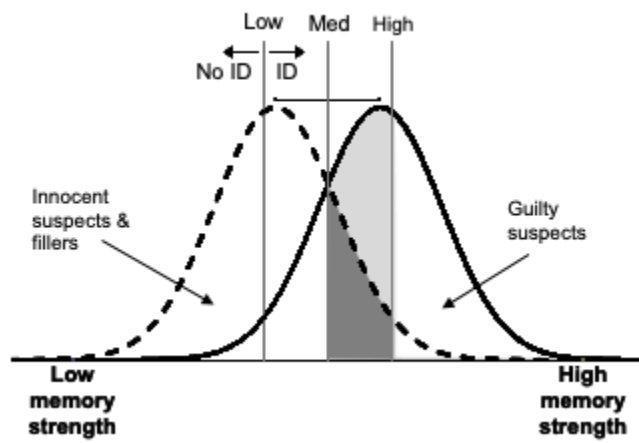
A signal-detection theory account holds that recognition decisions and confidence emerge simultaneously and are based on the same information. According to a signal-detection interpretation, each face in a lineup generates some memory signal. Over many witnesses viewing a fair lineup, these memory signals can be depicted as two distributions of memory signals. As shown in Figure 2, the guilty suspect distribution sits further right on the memory strength axis than the innocent suspect and filler distribution, because witnesses have seen the guilty suspect before, but have not seen the innocent suspect and fillers before. The distance between the two distributions is how well witnesses can tell apart guilty suspects from innocent suspects and fillers. A decision criterion is placed on the memory strength axis and an identification is made when the memory strength of a face exceeds it. The overall correct identification rate, regardless of confidence, is shown by the proportion of the guilty suspect distribution that exceeds the decision

criterion (shaded light grey, Figure 2a), whereas the overall incorrect identification rate of innocent suspects or fillers is shown by the proportion of the innocent suspect/filler distribution that exceeds the decision criterion (shaded dark grey, Figure 2a). Each level of confidence has its own decision criterion (e.g., low, medium, high confidence). An identification made with high confidence, for example, is made when the memory strength of the face being identified exceeds the right-most decision criterion (see Figure 2c). When an identification is made with high confidence, a larger proportion of the guilty suspect distribution (shaded light grey) exceeds the decision criterion than does the innocent suspect/filler distribution (shaded dark grey). Therefore, high-confidence identifications will be highly accurate. When an identification is made with medium confidence, however, a relatively large proportion of both the guilty suspect distribution (shaded light grey) and the innocent suspect/filler distribution (shaded dark grey) exceeds the criterion (Figure 2b). Therefore, medium-confidence identifications will be more inaccurate than high-confidence identifications. Likewise, low-confidence identifications will be more inaccurate than medium-confidence identifications.

A



B



C

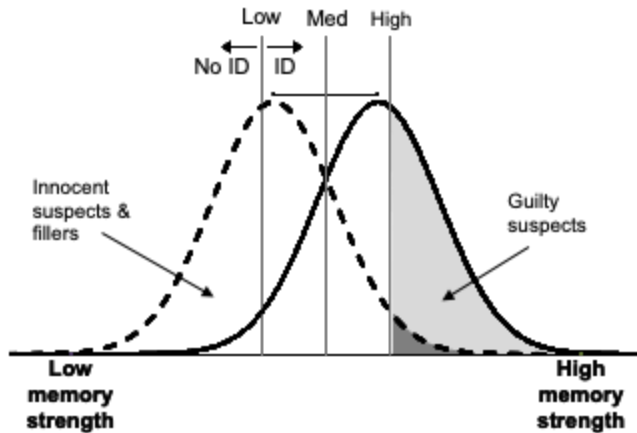


Figure 2. A signal-detection theory interpretation of a lineup task, illustrating identifications made with (a) any level of confidence (low, medium or high), (b) medium confidence and (c) high confidence. Note that some of the light grey shading is not visible because it is obscured by the overlapping dark grey shading.

Should the legal system rely on eyewitness confidence?

It is well known that eyewitness confidence can become inflated over time and with multiple memory retrieval attempts. Therefore, it is agreed that it is unwise to rely on eyewitness confidence that is provided only in the courtroom, several months or years after the incident (Wells et al., 2020). However, academics are not in agreement as to the extent to which legal systems should rely on eyewitness confidence collected on the *initial memory test*. Some advocate that collecting confidence is highly informative of accuracy and is therefore the most important factor to determine eyewitness accuracy, and more important than other estimator variables that might exist in a case (e.g., Wixted & Mickes, 2022).

Other academics argue that, while confidence can be informative, sometimes it is not, such as under biased memory retrieval conditions. Biased memory retrieval conditions include the police using an unfair lineup, where the suspect stands out from the fillers because he is the only person who matches the witness's description of the perpetrator (e.g., Colloff et al., 2016). Biased memory retrieval conditions also include the witness being asked suggestive questions (Loftus et al., 1978), or an abundance of closed questions (e.g., Lamb et al., 2007) that reduce the accuracy of the witness's memory reports. Because it can be difficult for investigators to know if these suboptimal conditions have prevailed, it might be imprudent to overly-rely on witness confidence (e.g., Berkowitz et al., 2020). Of course, the Houston Police Department study

described above suggests that confidence can be informative about identification accuracy for real eyewitnesses, though more research in other jurisdictions is needed.

System Variables

First, it is important to note that some system variable findings in the identification literature have been controversial because the analytical techniques that have been used to infer accuracy have sometimes been misleading (See Box 3).

Box 3: Measuring Eyewitness Accuracy and Simultaneous versus Sequential Lineup Presentation

To measure the probative value of a procedure much prior research has calculated the diagnosticity ratio (correct identification rate from target-present lineups/false alarm rate of innocent suspects from target-absent lineups). It is now clear that this measure is influenced by the propensity of witnesses to make an identification, called their response bias (Clark, 2012). When witnesses are more likely to make an identification, this means that they are more likely to correctly identify the culprit in the target-present lineup, but also more likely to identify an innocent suspect in a target-absent lineup. Put another way, the benefits (i.e., correct identifications) come at a cost (i.e., false identifications). A key goal of lineup procedures is to enhance witness discriminability—the ability to tell the difference between innocent and guilty suspects. A procedure that enhances discriminability yields more correct identifications of culprits in target-present lineups, without increasing incorrect identifications of innocent suspects.

Signal-detection theory and associated analyses allow discriminability to be measured independently from response bias. In 2012, Mickes et al. applied Receiver Operating Characteristic (or ROC) analysis to eyewitness data. Data are combined from different participants and plotted as a function of the level of confidence to produce a ROC curve. The higher the ROC curve, the better the procedure is at enabling eyewitnesses to tell the difference between innocent and guilty suspects. A measure of discriminability is now generally required to publish experiments on eyewitness identification.

Why does the analysis type matter? It matters because using the diagnosticity ratio can

lead to misleading conclusions regarding which eyewitness procedure maximizes eyewitness performance. As an example, there has been ongoing debate as to whether lineups should be conducted sequentially (i.e., presenting one person at a time), or simultaneously (i.e., presenting all of the people at the same time). Prior research found the diagnosticity ratio to be higher for sequential lineups (e.g., Steblay et al., 2001). As a result, the sequential lineup was once considered to be diagnostically superior to the simultaneous lineup, and implemented by several police agencies in the US, including North Carolina and Massachusetts. However, signal-detection theory and ROC analyses now indicate that the sequential procedure generally makes witnesses less willing to make an identification; it yields fewer false identifications, but it also yields fewer correct identifications (Clark, 2012). Put another way, the benefits come at a cost. When ROC analysis is conducted, simultaneous lineups often yield higher discriminability than sequential lineups (see Seale-Carlisle et al., 2019). Recent policy and procedure guidelines written by prominent scientists have not yet advocated for one procedure over the other, however (e.g., National Research Council, 2014; Wells et al., 2020). England and Wales typically use sequential presentation of video images.

Blind Administration

Eyewitness scientists agree that the person who administers the lineup procedure and the witness themselves should not know (i.e., is blind to) the identity of the suspect in a lineup. Lineup administrators can—advertently or inadvertently—influence the behaviour of the witness and the interpretation of evidence. For example, the administrator may look at the witness when a suspect is being viewed, increasing the likelihood of a suspect identification regardless of whether they are innocent or guilty (Kovera & Evelo, 2017), or provide positive feedback to the witness about their decision, inflating witness confidence (Wells & Bradfield, 1998). Non-blind administrators also record witness behaviour differently. In one study, non-blind administrators were more likely to incorrectly report that witnesses who identified a filler had not made an identification, and witness confidence in a suspect identification was interpreted as being higher (Rodriguez & Berry, 2019).

Lineup Fillers

Scientists also agree that police officers should choose fillers who do not make the suspect stand out in the lineup based upon physical appearances or other factors such as clothing or background (Wells et al., 2020). When the suspect stands out, this increases the likelihood that the suspect will be identified, regardless of whether they are innocent or guilty (e.g., Lindsay & Wells, 1980), and also makes it more difficult for witnesses to tell the difference between innocent and guilty lineup members (e.g., Colloff et al., 2016). There is currently debate regarding how fair fillers improve lineup performance. Some suggest that fair fillers help witnesses to decide which facial features are useful for the identification decision, called *diagnostic feature detection* (e.g., Colloff et al., 2016; Wixted & Mickes, 2014). Others suggest that fair fillers siphon false positive identifications away from the innocent suspect more than they siphon correct identifications away from the guilty suspect, called *filler siphoning* (e.g., Smith et al., 2017; Smith et al., 2018). It is possible that both processes play a role.

Although scientists agree that fair lineup fillers should be used, there is not yet agreement about the best strategy for choosing fillers. There are two main strategies: selecting fillers that match the description provided by the witness (e.g., “White male, blonde hair, small nose”) or match the physical appearance of the suspect. There is a concern that match-to-description fillers might be too dissimilar to each other and therefore put an innocent suspect at risk of misidentification. Whereas match-to-appearance fillers can result in lineups containing highly-similar fillers (even clones), making the identification task impossible. The research has yielded inconsistent results about which method best supports accuracy, possibly because it has not been guided by a strong theoretical framework in which to make predictions and interpret results (Colloff et al., 2021). Nevertheless, a recent large-scale study showed that description-matched fillers enhanced the ability of witnesses to tell the difference between innocent and guilty suspects compared to appearance-matched fillers (Carlson et al., 2019).

Colloff et al. (2021) recently introduced a formal feature-matching model to make predictions about the influence of filler similarity on eyewitness discriminability. The model predicted that a two-step process would best optimize lineup performance. Step One is to create a pool of description-matched fillers, all of whom would be suitable for creating a fair lineup because they would be reasonably described in the same way the witness described the perpetrator. Step Two is to *maximize dissimilarity* by choosing from that pool the fillers who are the least similar to the suspect. The model predicted that this strategy would increase correct

identifications of culprits without affecting the false-alarm rate of innocent suspects, and therefore increase discriminability. Interestingly, Wells et al. predicted this pattern of results in 1993 without the use of a formal model, but much of the subsequent research has considered description-matched lineups only (Step One) without also maximizing dissimilarity (Step Two). Colloff et al. (2021) confirmed the model's predictions in a large-scale experiment testing over 10,000 participants, suggesting the two-step process for selecting fillers might optimize eyewitness performance. Further research is, however, required.

Lineup Modality

Scientists have also extensively investigated whether accuracy varies in relation to the modality of presentation. Recent research finds that accuracy for live lineups, wherein the lineup members are physically present before the witness, does not vary compared to video lineups, where each lineup member is shown from the shoulders up, moving their head from left to right and then back to face the camera (Fitzgerald et al., 2018). Researchers are also testing a new interactive lineup, in which witnesses can freely rotate the lineup faces to view the faces from multiple angles (Colloff et al., 2022). This research has found that lineups presented in an interactive format improve accuracy compared to photo lineups where each lineup member is shown as a static image facing the camera.

Lineups versus Showups

Lineups embed the suspect among known-innocent fillers, whereas showups do not use fillers and instead present the suspect alone. In a showup, the witness is asked whether or not the police suspect is the culprit. Scientists agree that lineups produce better outcomes than do showups (e.g., Wells et al., 2020). Experiments comparing lineups with showups have found that lineups better enable witnesses to distinguish between innocent and guilty suspects (e.g., Neuschatz et al., 2016). Incorrect positive identifications always land on the innocent suspect in a showup. But fair fillers in lineups offer some protection against misidentifications of the innocent suspect (called *filler siphoning*; e.g., Wells, 2001). Fair fillers may also help witnesses to decide which facial features are useful for the identification decision, called *diagnostic feature detection* (e.g., Colloff & Wixted, 2020). In practice, showups are often conducted in the field during which search-and-detain operations are actively unfolding. After a crime has taken place,

and a suspect who fits the witness's description has been found in the vicinity, the choice is often between conducting a showup or setting the detained person free. Therefore, showups are still used in practice.

American Psychology-Law Society 2020 White Paper

Experts in the American Psychology-Law Society published nine recommendations to advance the reliability and integrity of eyewitness evidence (Wells et al., 2020):

1. ***Prelineup Interview.*** As soon as possible after a crime, the witness should be interviewed to collect important information (e.g., description of the culprit, viewing conditions).
2. ***Evidence-Based Suspicion.*** There should be good reason to suspect that an individual is guilty before an identification parade is conducted to minimize the number of false identifications of innocent suspects.
3. ***Double-Blind.*** The administrator and the witness should not know who the suspect is in the lineup, or methods should be employed to prevent an administrator from inadvertently influencing the witness.
4. ***Lineup Fillers.*** A lineup should contain one suspect and at least five fillers who do not make the suspect stand out.
5. ***Prelineup Instructions.*** A witness should be told information to prevent them from being predisposed to identify someone, including that the culprit may not be in the lineup, and they are able to respond, "I don't know".
6. ***Confidence.*** Immediately after making a decision, witnesses should be asked how confident they are because this information is a useful cue to witness accuracy.
7. ***Video-Recording the identification procedure.*** To accurately record the conditions of the procedure and the witness response.
8. ***Avoid Repeated Identification Procedures with the Same Witness and Suspect.***
Earlier identification tests can contaminate a later identification test including the same suspect, resulting in memory errors.
9. ***Showups.*** Use lineups rather than showups if possible. If using showups then employ procedural safeguards, such as warning the witness that the suspect may not be the culprit.

Official Guidance

Eyewitness Statements and Testimony

In England and Wales, the Government has produced the Achieving Best Evidence (hereafter ABE) guidance for conducting police interviews with vulnerable and intimidated witnesses, as well as for victims of sexual offenses (Ministry of Justice, 2011). The ABE recommends adopting a phased approach to interviewing. The phases include establishing rapport and setting ground rules for the interview with the witness, asking the witness to provide a free narrative account, asking the witness focused questions pertaining to the incident, and closing the interview. The guidance also emphasizes the importance of pre-interview planning to ensure witness appropriate questions are asked to obtain relevant and necessary information without causing undue distress to the interviewee.

In the first phase of the interview, the interviewer should first endeavour to establish a strong rapport with the interviewee to improve the quality and quantity of information elicited. The ‘ground rules’ of the interview should also be explained to ensure the witness has clear expectations and feels in control of the interview process. In the second phase, the witness is asked to provide a narrative account of the incident. The ABE instructs interviewers to engage in active listening, not interrupt the witness, and use prompts only when necessary to gather further information. In the next phase, the interviewer asks specific questions to follow-up on the information that the witness provided during their free narrative account for purposes of gathering more detailed information where possible. Witnesses should be told that it is acceptable for them to state that ‘I don’t know’ or ‘I don’t understand’. In the final phase, the interviewer might briefly summarise the information collected and provide an opportunity for the witness to add further information. There is a significant literature on both the ABE, as well as approaches that may be used for interviewing witnesses (see for examples: Fisher & Geiselman, 1992, Memon et al., 2010).

In the US, there are also well-established best practice guidelines for interviewing vulnerable witnesses, such as children (e.g., National Institute of Child Health and Human Development (NICHD) Structured Interview Protocol for Children, see Orbach et al., 2000). Internationally, guidance has been developed for interviews with rape victims (Flowe et al., 2021), for critical incidents such as terrorism (e.g., Smith & Milne, 2018), for interviewing

witnesses and victims for purposes of documenting and investigating sexual violence in conflict (FCDO, 2017), as well as good practice for interviewing suspects, such as the Mendez Principles which have preventing torture as their primary aim (IRCT, 2022).

Eyewitness Identification

In England and Wales, identification procedures are governed by Code D of the Police and Criminal Evidence Act (PACE) (1984) Codes of Practice, which was last updated in 2017. If a formal identification procedure is conducted, a video parade containing moving images must be offered unless it is not practicable, or a “live” in-person parade is considered more suitable. In a video parade, each person is shown sequentially. PACE also states that lineups should contain a minimum of eight fillers and one suspect. The fillers must “resemble the suspect in age, general appearance and position in life”. In 2017, a note was added to clarify that the requirement is not that the lineup members be identical or extremely similar. Witnesses should be told that the person they saw may not be present and view the entire lineup at least twice. The lineup administrator should not be involved in the investigation of the case and anything the witness says should be written down.

In the US identification procedures are regulated at the state level or in local jurisdictions, so there is no single code of practice. In 2017, however, the US Deputy Attorney General issued a policy which applies to all agencies in the Department of Justice (The U.S. Department of Justice, 2017). It outlines that the most common identification procedure uses photographs. The photo array should include the suspect and at least five fillers. The fillers should: “generally fit the witness's description of the perpetrator, including such characteristics as gender, race, skin color, facial hair, age, and distinctive physical features. They should be sufficiently similar so that a suspect's photograph does not stand out, but not so similar that a person who knew the suspect would find it difficult to distinguish him or her.” (p. 3). Witnesses should be told that the person they saw may or may not be present and are able to make an identification at any time and view the images for as long as they want. The procedure should, ideally, be conducted so that the administrator is “blind”. The administrator should ask the witness to state in his or her own words how confident they are in their identification. The procedure should be video or audio recorded, or a written statement of what the witness has said verbatim should be made.

Summary

- The legal system relies on information and identifications made by witnesses in investigating and prosecuting cases
- The accuracy and completeness of information witnesses provide can vary in relation to estimator and system variables
- The analytic approach used to examine the relationship between confidence and accuracy is important in ascertaining witness reliability
- Witness confidence in the accuracy of their statements, testimony, and identifications at the time of the initial test is a strong predictor of likely accuracy
- Psychological science has led to guidance for interviewing and identification procedures

Essay/Discussion Questions

1. “The Criminal Justice System should rely on eyewitness confidence to determine the accuracy of eyewitness evidence.” Discuss.
2. Describe two estimator variables and two system variables that might influence the accuracy of eyewitness evidence.
3. Critically evaluate current official guidance for collecting eyewitness evidence using findings from psychological science research.
4. “Witnesses who were alcohol intoxicated compared to sober at the time of the crime will not remember it as well.” Discuss.

Annotated Reading List

Loftus, E. F. (1996). *Eyewitness testimony* (2nd edn). Harvard: Harvard University Press.

This book, written by one of the best-known eyewitness researchers, provides a useful introduction to many of the topics outlined in this chapter.

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36(12), 1546–1557.

The highly influential article that first distinguished between estimator and system variables and their impact on eyewitness memory.

Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, 44(1), 3–36.

<https://doi.org/10.1037/lhb0000359>

This is a seminal American Psychology-Law Society Scientific Review paper, containing procedural recommendations for eyewitness identification globally, written by 6 prominent Professorial eyewitness researchers. It was the product of an extensive, multistep vetting process designed to ensure that it represents the best research, analysis, and recommendations to guide police policy.

Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science*, 13(3), 324-335.

<https://doi.org/10.1177/1745691617734878>

The authors make the case that eyewitness memory is not inherently unreliable, and this paper was published alongside commentaries from other leading eyewitness memory researchers with alternate views.

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(1), 10-65. <https://doi.org/10.1177/1529100616686966>

This article provides a comprehensive review of the literature examining the relationship between confidence and eyewitness accuracy, and provides an overview of legal considerations.

References

- Assefi, S. L., & Garry, M. (2003). Absolut® memory distortions: Alcohol placebos influence the misinformation effect. *Psychological Science*, 14(1), 77-80. <https://doi.org/10.1111/1467-9280.01422>
- Bartlett, J. C., & Fulton, A. (1991). Familiarity and recognition of faces in old age. *Memory & Cognition*, 19(3), 229-238. <https://doi.org/10.3758/BF03211147>

- Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, 25(5), 475-491. <https://doi.org/10.1023/A:1012840831846>
- Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(1), 139-145. <https://doi.org/10.1037/0278-7393.9.1.139>
- Berkowitz, S. R., Garrett, B. L., Fenn, K. M., & Loftus, E. F. (2022). Convicting with confidence? Why we should not over-rely on eyewitness confidence. *Memory*, 30(1), 10-15. <https://doi.org/10.1080/09658211.2020.1849308>
- Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, 3(2), 77-88. <https://doi.org/10.1016/j.jarmac.2014.03.005>
- Braun, K. A., & Loftus, E. F. (1998). Advertising's misinformation effect. *Applied Cognitive Psychology*, 12(6), 569-591. [https://doi.org/10.1002/\(SICI\)1099-0720\(199812\)12:6<569::AID-ACP539>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1099-0720(199812)12:6<569::AID-ACP539>3.0.CO;2-E)
- Brewer, N., & Day, K. (2005). The confidence-accuracy and decision latency-accuracy relationships in children's eyewitness identification. *Psychiatry, Psychology and Law*, 12(1), 119-128. <https://doi.org/10.1375/pplt.2005.12.1.119>
- Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence-accuracy relationship in eyewitness identification: The effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied*, 8(1), 44-56. <https://doi.org/10.1037/1076-898X.8.1.44>
- Carlson, C. A., Jones, A. R., Whittington, J. E., Lockamyeir, R. F., Carlson, M. A., & Wooten, A. R. (2019). Lineup fairness: Propitious heterogeneity and the diagnostic feature-detection hypothesis. *Cognitive Research: Principles and Implications*, 4(1), 1-16. <https://doi.org/10.1186/s41235-019-0172-5>
- Chandler, C. C. (1991). How memory for an event is influenced by related events: Interference in modified recognition tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(1), 115-125. <https://doi.org/10.1037/0278-7393.17.1.115>
- Chung, M. S., & Thomson, D. M. (1995). Development of face recognition. *British Journal of Psychology*, 86(1), 55-87. <https://doi.org/10.1111/j.2044-8295.1995.tb02546.x>
- Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science*, 7(3), 238-259. <https://doi.org/10.1177/1745691612439584>
- Clifasefi, S. L., Takarangi, M. K., & Bergman, J. S. (2006). Blind drunk: The effects of alcohol on inattentional blindness. *Applied Cognitive Psychology*, 20(5), 697-704. <https://doi.org/10.1002/acp.1222>
- Colloff, M. F., Flowe, H. D., Smith, H. M. J., Seale-Carlisle, T. M., Meissner, C. A., Rockey, J. C., Pande, B., Kujur, P., Parveen, N., Chandel, P., Singh, M. M., Pradhan, S., & Parganiha, A. (2022). Active exploration of faces in police lineups increases discrimination accuracy. *American Psychologist*, 77(2), 196-220. <https://doi.org/10.1037/amp0000832>
- Colloff, M. F., Wade, K. A., & Strange, D. (2016). Unfair lineups make witnesses more likely to confuse innocent and guilty suspects. *Psychological Science*, 27(9), 1227-1239. <https://doi.org/10.1177/0956797616655789>

- Colloff, M. F., Wade, K. A., Wixted, J. T., & Maylor, E. A. (2017). A signal-detection analysis of eyewitness identification across the adult lifespan. *Psychology and Aging*, 32(3), 243–258. <https://doi.org/10.1037/pag0000168>
- Colloff, M. F., Wilson, B. M., Seale-Carlisle, T. M., & Wixted, J. T. (2021). Optimizing the selection of fillers in police lineups. *Proceedings of the National Academy of Sciences*, 118(8), e2017292118. <https://doi.org/10.1073/pnas.2017292118>
- Colloff, M. F., & Wixted, J. T. (2020). Why are lineups better than showups? A test of the filler siphoning and enhanced discriminability accounts. *Journal of Experimental Psychology: Applied*, 26(1), 124–143. <https://doi.org/10.1037/xap0000218>
- Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, 111(2), 219–247. <https://doi.org/10.1016/j.cognition.2009.02.004>
- Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime and Law*, 6(2), 81–97. <https://doi.org/10.1080/10683160008410834>
- Devlin, Lord P. 1976. *Report to the Secretary of State for the Home Department on the Departmental Committee on Evidence of Identification in Criminal Cases*, London: HMSO.
- Dodson, C. S., & Dobolyi, D. G. (2016). Confidence and eyewitness identifications: The cross-race effect, decision time and accuracy. *Applied Cognitive Psychology*, 30(1), 113–125. <https://doi.org/10.1002/acp.3178>
- Dodson, C. S., & Krueger, L. E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin & Review*, 13(5), 770–775. <https://doi.org/10.3758/BF03193995>
- Dunlevy, J. R., & Cherryman, J. (2013). Target-absent eyewitness identification line-ups: Why do children like to choose. *Psychiatry, Psychology and Law*, 20(2), 284–293. <https://doi.org/10.1080/13218719.2012.671584>
- Ebbesen, E. B., & Rienick, C. B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology*, 83(5), 745–762.
- Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.
- Evans, J. R., Schreiber Compo, N., & Russano, M. B. (2009). Intoxicated witnesses and suspects: Procedures and prevalence according to law enforcement. *Psychology, Public Policy, and Law*, 15(3), 194–221. <https://doi.org/10.1037/a0016837>
- 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., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, 15(3), 177–185.
- Fitzgerald, R. J., & Price, H. L. (2015). Eyewitness identification across the life span: A meta-analysis of age differences. *Psychological Bulletin*, 141(6), 1228–1265. <https://doi.org/10.1037/bul0000013>
- Fitzgerald, R. J., Price, H. L., & Valentine, T. (2018). Eyewitness identification: Live, photo, and video lineups. *Psychology, Public Policy, and Law*, 24(3), 307–325. <https://doi.org/10.1037/law0000164>
- Flowe, H. D., Humphries, J. E., Takarangi, M. K., Zelek, K., Karoğlu, N., Gabbert, F., & Hope, L. (2019). An experimental examination of the effects of alcohol consumption and exposure to

- misleading postevent information on remembering a hypothetical rape scenario. *Applied Cognitive Psychology*, 33(3), 393-413. <https://doi.org/10.1002/acp.3531>
- Flowe, H. D., Takarangi, M. K., Humphries, J. E., & Wright, D. S. (2016). Alcohol and remembering a hypothetical sexual assault: Can people who were under the influence of alcohol during the event provide accurate testimony?. *Memory*, 24(8), 1042-1061. <https://doi.org/10.1080/09658211.2015.1064536>
- Foreign and Commonwealth Office (2017). *International Protocol on the Documentation and Investigation of Sexual Violence in Conflict* (second ed.), Foreign and Commonwealth Office, London.
- Gabbert, F., Hope, L., & Fisher, R. P. (2009). Protecting eyewitness evidence: Examining the efficacy of a self-administered interview tool. *Law and Human Behavior*, 33(4), 298-307. <https://doi.org/10.1007/s10979-008-9146-8>
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event?. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 17(5), 533-543. <https://doi.org/10.1002/acp.885>
- Goldsmith, M., Koriati, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language*, 52(4), 505-525. <https://doi.org/10.1016/j.jml.2005.01.010>
- Goodwin, D. W., Crane, J. B., & Guze, S. B. (1969a). Alcoholic "blackouts": A review and clinical study of 100 alcoholics. *American Journal of Psychiatry*, 126(2), 191-198. <https://doi.org/10.1176/ajp.126.2.191>
- Goodwin, D., Crane, J., & Guze, S. (1969b). Phenomenological Aspects of the Alcoholic "Blackout". *British Journal of Psychiatry*, 115(526), 1033-1038. <https://doi.org/10.1192/bjp.115.526.1033>
- Greene, E., & Fraser, S. C. (2002). Observation distance and recognition of photographs of celebrities' faces. *Perceptual & Motor Skills*, 95(2), 637-651. <https://doi.org/10.2466/pms.2002.95.2.637>
- Hope, L., Gabbert, F., Fisher, R. P., & Jamieson, K. (2014). Protecting and enhancing eyewitness memory: The impact of an initial recall attempt on performance in an investigative interview. *Applied Cognitive Psychology*, 28(3), 304-313. <https://doi.org/10.1002/acp.2984>
- Hope, L., Mullis, R., & Gabbert, F. (2013). Who? What? When? Using a timeline technique to facilitate recall of a complex event. *Journal of Applied Research in Memory and Cognition*, 2(1), 20-24. <https://doi.org/10.1016/j.jarmac.2013.01.002>
- Humphries, J. E., & Flowe, H. D. (2015). Receiver operating characteristic analysis of age-related changes in lineup performance. *Journal of Experimental Child Psychology*, 132, 189-204. <https://doi.org/10.1016/j.jecp.2014.12.009>
- Innocence Project. (2019). DNA exonerations in the United States. <https://www.innocenceproject.org/dna-exonerations-in-the-united-states/>
- International Rehabilitation Council for Torture Victims. (2022, January 21). *IRCT: Mendez Principles a 'Huge Step Forward in Prevention of Torture'*. <https://www.irct.org/media-and-resources/latest-news/article/1114>
- Jores, T., Colloff, M. F., Kloft, L., Smailes, H., & Flowe, H. D. (2019). A meta-analysis of the effects of acute alcohol intoxication on witness recall. *Applied Cognitive Psychology*, 33(3), 334-343. <https://doi.org/10.1002/acp.3533>

- 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(5), 1304–1316. <https://doi.org/10.1037/0278-7393.22.5.1304>
- Keast, A., Brewer, N., & Wells, G. L. (2007). Children’s metacognitive judgments in an eyewitness identification task. *Journal of Experimental Child Psychology*, 97(4), 286–314. <https://doi.org/10.1016/j.jecp.2007.01.007>
- Kelly, L., Lovett, J. & Regan, L. (2005) *A Gap or a Chasm? Attrition in Reported Rape Cases*, Home Office Research Study 293, London: Home Office. Available online at: <http://www.homeoffice.gov.uk/rds/pdfs05/hors293.pdf>
- Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, 123(3), 297–315. <https://doi.org/10.1037/0096-3445.123.3.297>
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, 103(3), 490–517. <https://doi.org/10.1037/0033-295X.103.3.490>
- Kovera, M. B., & Evelo, A. J. (2017). The case for double-blind lineup administration. *Psychology, Public Policy, and Law*, 23(4), 421–437. <http://dx.doi.org/10.1037/law0000139>
- 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(9), 1117–1130. <https://doi.org/10.1002/acp.1318>
- Lindsay, D. S., & Johnson, M. K. (1987). Reality monitoring and suggestibility: Children’s ability to discriminate among memories from different sources. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children’s eyewitness memory* (pp. 92–121). New York: Springer.
- Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, 9(3), 215–218. <https://doi.org/10.1111/1467-9280.00041>
- Lindsay, R. C. L., Semmler, C., Weber, N., Brewer, N., & Lindsay, M. R. (2008). How variations in distance affect eyewitness reports and identification accuracy. *Law and Human Behavior*, 32(6), 526–535. <https://doi.org/10.1007/s10979-008-9128-x>
- Lindsay, R. C. L., & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior*, 4(4), 303–313. <https://doi.org/10.1007/BF01040622>
- Lockamy, R. F., Carlson, C. A., Jones, A. R., Carlson, M. A., & Weatherford, D. R. (2020). The effect of viewing distance on empirical discriminability and the confidence–accuracy relationship for eyewitness identification. *Applied Cognitive Psychology*, 34(5), 1047–1060. <https://doi.org/10.1002/acp.3683>
- Loftus, E. F. (1979). The malleability of human memory: Information introduced after we view an incident can transform memory. *American Scientist*, 67(3), 312–320.
- Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin & Review*, 12(1), 43–65. <http://dx.doi.org/10.3758/BF03196348>
- Loftus, E. F., & Hoffman, H. G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General*, 118(1), 100–104. <https://doi.org/10.1037/0096-3445.118.1.100>

- 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(1), 19–31. <https://doi.org/10.1037/0278-7393.4.1.19>
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585–589. [https://doi.org/10.1016/S0022-5371\(74\)80011-3](https://doi.org/10.1016/S0022-5371(74)80011-3)
- Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, 25(12), 720–725. <https://doi.org/10.3928/0048-5713-19951201-07>
- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, 13(4), 330–334. <https://doi.org/10.1037/h0028434>
- McCloskey, M., & Zaragoza, M. (1985a). Misleading post event information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, 114(1), 1–16.
- McCloskey, M., & Zaragoza, M. (1985b). Postevent information and memory: Reply to Loftus, Schooler, and Wagenaar. *Journal of Experimental Psychology: General*, 114(3), 381–387. <https://doi.org/10.1037/0096-3445.114.3.381>
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3–35. <https://doi.org/10.1037/1076-8971.7.1.3>
- Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology*, 19(5), 545–567. <https://doi.org/10.1002/acp.1097>
- Memon, A., & Higham, P. A. (1999). A review of the cognitive interview. *Psychology, Crime and Law*, 5(1-2), 177–196. <https://doi.org/10.1080/10683169908415000>
- Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition*, 30(8), 1219–1227. <https://doi.org/10.3758/BF03213404>
- Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, 16(4), 340–372. <https://doi.org/10.1037/a0020518>
- 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(2), 93–102. <https://doi.org/10.1016/j.jarmac.2015.01.003>
- Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied*, 18(4), 361–376. <https://doi.org/10.1037/a0030609>
- Ministry of Justice. (2011). *Achieving best evidence in criminal proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures*. London: Ministry of Justice.
- National Research Council (2014). *Identifying the Culprit: Assessing Eyewitness Identification*. Washington, DC: The National Academies Press.
- Neuschatz, J. S., Preston, E. L., Burkett, A. D., Togliola, M. P., Lampinen, J. M., Neuschatz, J. S., ... & Goodsell, C. A. (2005). The effects of post-identification feedback and age on retrospective

- eyewitness memory. *Applied Cognitive Psychology*, 19(4), 435-453.
<https://doi.org/10.1002/acp.1084>
- Neuschatz, J. S., Wetmore, S. A., Key, K. N., Cash, D. K., Gronlund, S. D., & Goodsell, C. A. (2016). A comprehensive evaluation of showups. In M. K. Miller & B. H. Bornstein (Eds.), *Advances in psychology and law* (pp. 43–69). Springer International Publishing.
https://doi.org/10.1007/978-3-319-29406-3_2
- Nguyen, T. B., Pezdek, K., & Wixted, J. T. (2017). Evidence for a confidence-accuracy relationship in memory for same- and cross-race faces. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 70(12), 2518–2534. <https://doi.org/10.1080/17470218.2016.1246578>
- Nyman, T. J., Antfolk, J., Lampinen, J. M., Tuomisto, M., Kaakinen, J. K., Korkman, J., & Santtila, P. (2019). A stab in the dark: The distance threshold of target identification in low light. *Cogent Psychology*, 6(1), 527–541. <https://doi.org/10.1080/23311908.2019.1632047>
- Nyman, T. J., Lampinen, J. M., Antfolk, J., Korkman, J., & Santtila, P. (2019). The distance threshold of reliable eyewitness identification. *Law and Human Behavior*, 43(6), 527–541. <https://doi.org/10.1037/lhb0000342>
- Office of National Statistics. (2011/2012). *Focus on violent crime & sexual offences*. Retrieved from http://www.ons.gov.uk/ons/dcp171778_298904.pdf
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2000). Interviewing at the scene of a crime: Effects on children's recall of alleged abuse. *Legal and Criminological Psychology*, 5(1), 135–147. <https://doi.org/10.1348/135532500167930>
- Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, 19(1), 55–71.
<https://doi.org/10.1037/a0031602>
- Palmer, F. T., Flowe, H. D., Takarangi, M. K. T., & Humphries, J. E. (2013). Intoxicated witnesses and suspects: An archival analysis of their involvement in criminal case processing. *Law and Human Behavior*, 37(1), 54–59. <https://doi.org/10.1037/lhb0000010>
- Police and Criminal Evidence Act (1984). *Codes of Practice, Code D* (2017) Retrieved from: <https://www.gov.uk/guidance/police-and-criminal-evidence-act-1984-pace-codes-of-practice>
- Pozzulo, J. D., & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, 22(5), 549-570.
<https://doi.org/10.1023/A:1025739514042>
- R v Turnbull and others (1976). 63 Cr App R 132.
- Read, J. D., Lindsay, D. S., & Nichols, T. (1998). The relation between confidence and accuracy in eyewitness identification studies: Is the conclusion changing? In C. P. Thomson, D. Bruce, J. D. Read, D. Hermann, D. Payne, & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 107–130). Mahwah, NJ: Lawrence Erlbaum.
- Rodriguez, D. N., & Berry, M. A. (2014). The effect of line-up administrator blindness on the recording of eyewitness identification decisions. *Legal and Criminological Psychology*, 19(1), 69–79. <http://dx.doi.org/10.1111/j.2044-8333.2012.02058.x>
- Sauer, J., Brewer, N., Zuck, T., & Weber, N. (2010). The effect of retention interval on the confidence-accuracy relationship for eyewitness identification. *Law and Human Behavior*, 34(4), 337–347. <https://doi.org/10.1007/s10979-009-9192-x>

- Seale-Carlisle, T. M., Wetmore, S. A., Flowe, H. D., & Mickes, L. (2019). Designing police lineups to maximize memory performance. *Journal of Experimental Psychology: Applied*, 25(3), 410–430. <https://doi.org/10.1037/xap0000222>
- Searcy, J. H., Bartlett, J. C., & Memon, A. (1999). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Memory & Cognition*, 27(3), 538-552. <https://doi.org/10.3758/BF03211547>
- Searcy, J., Bartlett, J. C., & Memon, A. (2000). Influence of post-event narratives, line-up conditions and individual differences on false identification by young and older eyewitnesses. *Legal and Criminological Psychology*, 5(2), 219-235. <https://doi.org/10.1348/135532500168100>
- Semmler, C., Dunn, J., Mickes, L., & Wixted, J. T. (2018). The role of estimator variables in eyewitness identification. *Journal of Experimental Psychology: Applied*, 24(3), 400–415. <https://doi.org/10.1037/xap0000157>
- Smith, A. M., Wells, G. L., Lindsay, R. C. L., & Penrod, S. D. (2017). Fair lineups are better than biased lineups and showups, but not because they increase underlying discriminability. *Law and Human Behavior*, 41(2), 127–145. <https://doi.org/10.1037/lhb0000219>
- Smith, A. M., Wells, G. L., Smalarz, L., & Lampinen, J. M. (2018). Increasing the similarity of lineup fillers to the suspect improves the applied value of lineups without improving memory performance: Commentary on Colloff, Wade, and Strange (2016). *Psychological Science*, 29(9), 1548-1551. <https://doi.org/10.1177/0956797617698528>
- Smith, K., & Milne, B. (2018). Witness interview strategy for critical incidents (WISCI). *The Journal of Forensic Practice*, 20(4), 268-278. <https://doi.org/10.1108/JFP-03-2018-0007>
- Sodian, B., Thoermer, C., Kristen, S., & Perst, H. (2012). Meta-cognition in infants and young children. In M. J. Beran, J. L. Brandl, J. Perner & J. Proust (Eds.), *Foundations of metacognition*. Oxford, England: Oxford University Press.
- Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, 7(1), 36–97. <https://doi.org/10.1037/1076-8971.7.1.36>
- Sporer, S. L., & Martschuk, N. (2014). The reliability of eyewitness identifications by the elderly: An evidence-based review. In M. P. Toglia, D. F. Ross, J. Pozzulo, & E. Pica (Eds.), *The elderly eyewitness in court* (pp. 3–37). New York, NY: Psychology Press.
- Stebly, N., Dysart, J., Fulero, S., & Lindsay, R. C. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, 25(5), 459-473. <https://doi.org/10.1023/A:1012888715007>
- Stebly, N. K., Wells, G. L., & Douglass, A. B. (2014). The eyewitness post identification feedback effect 15 years later: Theoretical and policy implications. *Psychology, Public Policy, and Law*, 20(1), 1–18. <https://doi.org/10.1037/law0000001>
- Tuckey, M. R., & Brewer, N. (2003). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied*, 9(2), 101–118. <https://doi.org/10.1037/1076-898X.9.2.101>
- United States Department of Justice. (2017). *Eyewitness identification: Procedures for conducting photo arrays*. <https://www.justice.gov/file/923201/download>
- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, 17(8), 969-993. <https://doi.org/10.1002/acp.939>
- Van Koppen, P. J., & Lochun, S. K. (1997). Portraying perpetrators; the validity of offender descriptions by witnesses. *Law and Human Behavior*, 21(6), 661-685. <https://doi.org/10.1023/A:1024812831576>

- Vredeveltdt, A., Hitch, G. J., & Baddeley, A. D. (2011). Eyeclosure helps memory by reducing cognitive load and enhancing visualisation. *Memory & Cognition*, 39(7), 1253-1263. <https://doi.org/10.3758/s13421-011-0098-8>
- Wagstaff, G. F., MacVeigh, J., Boston, R., Scott, L., Brunas-Wagstaff, J., & Cole, J. (2003). Can laboratory findings on eyewitness testimony be generalized to the real world? An archival analysis of the influence of violence, weapon presence, and age on eyewitness accuracy. *The Journal of Psychology*, 137(1), 17-28. <https://doi.org/10.1080/00223980309600596>
- Weber, N., & Brewer, N. (2008). Eyewitness recall: Regulation of grain size and the role of confidence. *Journal of Experimental Psychology: Applied*, 14(1), 50-60. <https://doi.org/10.1037/1076-898X.14.1.50>
- Weber, N., & Perfect, T. (2013). Why telling a witness that it's OK to say they don't know is good for justice. *Jury Expert*, 25(3), 36-42.
- Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36(12), 1546-1557. <https://doi.org/10.1037/0022-3514.36.12.1546>
- Wells, G. L. (2001). Eyewitness lineups: Data, theory, and policy. *Psychology, Public Policy, and Law*, 7, 791-801. <http://dx.doi.org/10.1037/1076-8971.7.4.791>
- 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(3), 360-376. <https://doi.org/10.1037/0021-9010.83.3.360>
- Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, 44(1), 3-36. <https://doi.org/10.1037/lhb0000359>
- Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, 78(5), 835-844. <https://doi.org/10.1037/0021-9010.78.5.835>
- White, A. M. (2003). What happened? Alcohol, memory blackouts, and the brain. *Alcohol Research & Health*, 27(2), 186- 196.
- Wilcock, R. A., Bull, R., & Vrij, A. (2005). Aiding the performance of older eyewitnesses: Enhanced non-biased line-up instructions and line-up presentation. *Psychiatry, Psychology and Law*, 12(1), 129-140. <https://doi.org/10.1375/pplt.2005.12.1.129>
- Winsor, A. A., Flowe, H. D., Seale-Carlisle, T. M., Killeen, I. M., Hett, D., Jores, T., Ingham, M., Lee, B. P., Stevens, L. M., & Colloff, M. F. (2021). Child witness expressions of certainty are informative. *Journal of Experimental Psychology: General*, 150(11), 2387-2407. <https://doi.org/10.1037/xge0001049>
- Wixted, J. T. (2022). The enigma of forgetting. *Proceedings of the National Academy of Sciences*, 119(12), e2201332119.
- Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, 2(6), 409-415. <https://doi.org/10.1111/j.1467-9280.1991.tb00175.x>
- Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, 121(2), 262-276. <https://doi.org/10.1037/a0035940>
- Wixted, J. T., & Mickes, L. (2022). Eyewitness memory is reliable, but the criminal justice system is not. *Memory*, 30, 67-72. <https://doi.org/10.1080/09658211.2021.1974485>

- 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(2), 304-309. <https://doi.org/10.1073/pnas.1516814112>
- Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science*, 13(3), 324-335. <https://doi.org/10.1177/1745691617734878>
- 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(2), 192-203. <https://doi.org/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(1), 10-65. <https://doi.org/10.1177/1529100616686966>
- Wright, D. B., & Loftus, E. F. (1998). How misinformation alters memories. *Journal of Experimental Child Psychology*, 71(2), 155–164. <https://doi.org/10.1006/jecp.1998.2467>
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3), 441–517. <https://doi.org/10.1006/jmla.2002.2864>
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71(2), 291–301. <https://doi.org/10.1037/0021-9010.71.2.291>
- Zaragoza, M. S., McCloskey, M., & Jamis, M. (1987). Misleading postevent information and recall of the original event: Further evidence against the memory impairment hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(1), 36–44. <https://doi.org/10.1037/0278-7393.13.1.36>