



Target Article

Eyewitness Identification in Its Social Context

Margaret Bull Kovera*

John Jay College and the Graduate Center, City University of New York, USA

Andrew J. Evelo

University of Waikato, New Zealand

Much of the literature on eyewitness identification neglects the social context in which identifications are made. As the number of cognitive psychologists conducting eyewitness research increased so did the use of signal detection theory and ROC analyses. With the resulting need for larger sample size, researchers moved toward conducting studies on internet platforms that allow for crowd-sourcing research participants. These methods make it next to impossible to ask research questions that explore the ways in which social interactions influence the identifications made by witnesses. Yet, the possibility of social context effects on witness memory are prevalent in applied contexts and research supports their existence. In addition, some eyewitness identifications may not be governed by memory at all. We argue that a consideration of social context effects is required to fully explore the reliability of witness identifications and propose a number of avenues for future research.

Keywords: Confidence, Eyewitness identification accuracy, Research methods, Signal detection theory, Social influence

General Audience Summary

Much of the current research on eyewitness identification focuses on the cognitive processes underlying witness identification and uses statistical analyses that require data from large numbers of participants to conduct properly. The need for so many participants led many eyewitness researchers to conduct their studies online, using crowd-sourcing methods to recruit participants and limiting the ability of researchers to ask research questions about the effects of social influence on identification accuracy. We argue that a consideration of social influence effects is necessary to fully understand the reliability of witness identifications.

In recent years, some scholars have argued that eyewitness identifications are more reliable than once thought, especially identifications that are made with high confidence (Brewin et al., 2020; Semmler et al., 2018; Wixted & Wells, 2017). Expert witnesses have even provided this opinion in court proceedings (e.g., *People of the State of New York v. Boone*, 2019). Given that this opinion is being offered in court, it is important to evaluate claims that eyewitness identifications are relatively reliable, especially when made with high confidence, and identify any deficiencies in the evidence used to support these claims. One deficiency is that much of the research investigating the factors that affect the reliability of

eyewitness identifications has been devoid of the psychosocial context in which these decisions are made in real cases.

Despite the role of social context in the accuracy of eyewitness identification, many of the studies that form the foundation of the argument that high confidence identifications can be trusted to be reliable, irrespective of the encoding conditions, were conducted recently, using pristine conditions that remove the possibility of social influence (e.g., Carlson et al., 2017; Lockamy et al., 2020; Pezdek et al., 2020). Some who initially advocated for the reliability of all high-confidence identifications have recognized that this accuracy-confidence relationship does not hold under all conditions

* Correspondence concerning this article should be addressed to Margaret Bull Kovera, Department of Psychology, John Jay College of Criminal Justice, City University of New York, NY, United States. Contact: mkovera@jjay.cuny.edu (M.B.K.)

(Wixted & Wells, 2017). Indeed, for the accuracy-confidence relationship to hold, eyewitness identifications must be collected under “pristine” conditions, which include a lineup administrator who does not know which lineup member is the suspect, instructions that warn that the culprit may not be present, and collection of confidence statement before any feedback about whether the witness has identified the suspect can be conveyed (Wixted & Wells, 2017). These pristine conditions constrain the effects of social influence on witnesses’ identification decisions, reflecting the importance of social context for the nature of the relationship between eyewitness accuracy and their confidence.

Yet surveys of police practice raise questions about whether these pristine conditions are frequently obtained in the real world (Police Executive Research Foundation, 2013). Of the departments responding to the survey, almost 40% of departments conducting live lineups did not collect a confidence statement from the witness when a positive identification was made. Over two-thirds of photo arrays and over 90% of live-lineups were conducted by administrators who were not blind to which lineup member was the suspect. Even though it is likely that the rates of blind administration are somewhat higher given changes in laws and practice since the survey was conducted (Kovera & Evelo, 2017), there remains ample room in actual police procedures for the social context to influence the decisions that witnesses make and the confidence that they report in the accuracy of those decisions.

In the remainder of this article, we review three lines of research that have demonstrated the importance of the social context in which identification decisions are made. We then argue that a number of recent trends in the study of eyewitness reliability have drawn attention away from the psychosocial context in which eyewitness identifications occur. These trends include an influx of cognitive psychologists into the area, bringing with them an emphasis on signal-detection theory (SDT) and analyses, and an increased use of crowd-sourced research participants to facilitate the data collection efforts that are required for these analyses. The number of participants required and the data collection platforms used to collect those data make it difficult to design and carry out research protocols that capture the social context in which identification decisions are typically made, limiting not only external validity (i.e., generalizability to other settings, including those in which identification procedures are more likely to occur) but also the types of questions that researchers will or can ask. We conclude with suggestions for research questions that acknowledge the importance of social context.

Social Influence Effects on Eyewitness Behavior

The argument that highly confident witnesses are accurate has reached courtrooms (*People of the State of New York v. Boone*, 2019), despite the clear need to qualify that conclusion with the limitation that the identification and confidence statements were collected under pristine conditions (Wixted & Wells, 2017). Many non-pristine practices should be avoided because they allow space for social context to influence witness decisions, attenuating the confidence-accuracy relationship. In

particular, there are substantial bodies of research demonstrating that interactions between lineup administrators and witnesses influence witnesses’ identification accuracy and their confidence statements. There are likely other sources of social influence that interfere with the confidence-accuracy relationship, including information gleaned from co-witnesses to an event. Next, we review the research demonstrating the social influence that both co-witnesses and administrators have on witnesses’ identification decisions and self-reported confidence.

Co-Witness Influence on Eyewitness Accuracy

Before witnesses ever attempt to make an identification, their memories for the culprit may become contaminated by their social interactions with co-witnesses to the same event (Wright et al., 2009). Misinformation from a co-witness about the appearance of a suspect can harm the accuracy of witnesses’ descriptions (Thorley & Kumar, 2017) as well as the accuracy of their identifications (Zajac et al., 2016; Zajac & Henderson, 2009). When witnesses heard from a confederate co-witness that the perpetrator of the crime that they had viewed was blue-eyed, when in fact he had brown eyes, the eyewitness participants were more likely to identify someone from a target-absent lineup consisting of blue-eyed lineup members than if they had not heard this misinformation (Experiments 1 and 3; Zajac & Henderson, 2009). This misinformation did not affect participants’ identifications from the same lineup when the researchers digitally altered the lineup members’ eyes to appear brown (Experiment 2). In another study, witnesses were more likely to misidentify a filler with a neck tattoo, especially at longer retention intervals, when a co-witness mistakenly reported that the perpetrator had a neck tattoo (Eisen, Gabbert et al., 2017). Misinformation from co-witnesses can even cause witnesses to claim that a person committed a crime when the witness did not see the crime that had been committed (Gabbert et al., 2003).

Many studies of co-witness influence have conveyed co-witness misinformation through a written description of what the co-witness had said (e.g., Thorley & Kumar, 2017). In real crime investigations, however, co-witness misinformation about the suspect is often conveyed during discussions between the witness and a co-witness. When discussing events in the context of a social encounter, (a) monitoring information sources is more difficult (Jack et al., 2014), perhaps because one has the added cognitive burden of generating what one will contribute to the conversation while listening to what the other person is saying; and (b) there will be conformity pressures present in the social setting (e.g., arriving upon a similar construal of the witnessed event will facilitate pleasant social interaction) that are likely absent from a situation in which one reads a report of what someone else said. Thus, it is not surprising that co-witness misinformation received through a discussion had a greater effect on witness memory than did misinformation received from a non-social source, like a newspaper or written statement (Gabbert et al., 2004; Paterson & Kemp, 2006). These findings highlight the importance of studying memory phenomena in the social settings in which

they occur because the social setting may alter the strength or the nature of a variable's influence on the identification of the culprit.

Information can be conveyed through witness behavior as well, even if that behavior does not contain information that would steer the witness toward identifying a particular lineup member. Instead, the information may alter witnesses' criteria for making a positive identification. For example, even if co-witnesses were prevented from talking with each other about the characteristics of the culprit, information gleaned from the identification behavior of a co-witness can influence witness identification choices. For example, witnesses who heard that their co-witness made an identification from the photo array were more likely to also make a positive identification from the array than were witnesses who heard their co-witness had rejected the array (Levett, 2013). In another study, witnesses were more likely to make a positive identification from a target-absent photo array when their co-witness made a fast versus slow identification decision (Experiment 1; Douglass et al., 2020). This effect of co-witness identification speed on witnesses' willingness to make a positive identification held even when it was clear that the witness had made a positive identification (Experiment 2) or had rejected the photo array (Experiment 3; Douglass et al., 2020).

This form of influence is known as informational influence in the social psychological literature, where it is often distinguished from normative influence (Deutsch & Gerard, 1955). Informational influence affects eyewitness identification by providing information—either regarding the suspect's identity or presence in the lineup—to the witness. The information itself can map onto SDT concepts, where additional information may affect assessments of familiarity (e.g., signal match to memory) or placement of the criterion (e.g., the likelihood the culprit is present in the lineup). Unlike informational influence, normative influence occurs when witnesses are swayed to make a particular choice not because they believe that choice is correct but because they fear the repercussions of failing to conform their behavior to what someone else desires.

Although some relatively early work acknowledged the relative contributions of informational and normative influence processes to co-witness effects on witness behaviors (Wright et al., 2009), few studies have tested the boundaries of normative influence in this context. Perhaps witnesses may not accept co-witnesses' versions of events as true but instead acquiesce to their versions to facilitate future social interactions with their co-witnesses or to cooperate with the police. In essence, the witnesses' memories are not changed by the misinformation provided by the co-witness, but their decisions are altered nonetheless, neither because of changes in discriminability nor criterion. Understanding the moderators of these social influence processes is important if we are to develop a full understanding of the variables affecting eyewitness behavior.

Administrator Influence on Eyewitness Accuracy

Police officers also may communicate information to witnesses that could alter the accuracy of their identification deci-

sions without altering the ability of witnesses to discriminate among guilty and innocent suspects or altering the conservativeness of their criterion for making an identification. Much like interrogators communicate details of the crime to innocent suspects during interrogations (Alceste et al., 2020), police officers could provide witnesses with details about a suspect's characteristics while interviewing them about what they saw, especially if the officers have pre-existing expectations about who might be likely to have committed a particular crime in a particular neighborhood (i.e., they have "usual suspects"). Probably more common, given the number of jurisdictions that continue to use single-blind lineup procedures (Kovera & Evelo, 2017), is a situation in which a police officer, who knows which lineup member is the suspect, administers the identification procedure to a witness and consciously or unconsciously steers the witness toward identifying the suspect, irrespective of the suspect's guilt (Kovera & Evelo, 2017, 2020).

In studies in which participants served as both lineup administrators and witnesses, non-blind administrators (i.e., administrators who knew which lineup member was the suspect) were more likely to smile when a witness looked at a photo of a suspect (Charman & Quiroz, 2016) or to point or draw attention to the suspect (Zimmerman et al., 2017) than were "blind" administrators who did not know. Fourteen percent of blind administrators asked the witness a question about the suspect's photo (rather than a filler's photo), a rate that is not significantly more than chance given a six-person photo array (Zimmerman et al., 2017). In contrast, 73% of non-blind administrators asked about the suspect's photo, clearly demonstrating that when administrators know which photo depicts the suspect, they will engage in behaviors that steer the witness toward an identification of the suspect.

Although a criterion shift can explain the changes in witness decisions when the administrator is merely steering the witness to make an identification (Clark et al., 2009), it cannot explain findings from studies in which administrators steer witnesses toward identifying the suspect (Clark et al., 2013; Evelo, 2020) or those comparing the different decisions made by witnesses administered by blind and non-blind administrators (Charman & Quiroz, 2016; Greathouse & Kovera, 2009; Zimmerman et al., 2017). In these studies, differential administrator behavior toward the suspect produced a reliable filler-to-suspect shift in witness identifications. The increases in identifications of the suspect that are obtained in single-blind lineup administrations are at the expense of filler identifications that would have been made if the administration had been double-blind; the rate of lineup rejections typically remains unchanged between single- and double-blind administrations (Charman & Quiroz, 2016; Greathouse & Kovera, 2009; Kovera & Evelo, 2017).

These findings are inconsistent with signal-detection models (Macmillan & Creelman, 2010). If discriminability was affected by whether the lineup administrator knew which lineup member was the suspect, then there should be conditions under which correct identifications increase and mistaken identifications decrease. Instead, non-blind administrators increase identifications of the suspect irrespective of whether the suspect

is guilty or innocent (Charman & Quiroz, 2016; Greathouse & Kovera, 2009). If administrator knowledge influenced criterion, then there would be changes in rejection rates across double- and single-blind conditions. In sum, nonblind administrators increase witness identifications of suspects through informational influence that is impermissibly suggestive rather than affecting discriminability or criterion. It is difficult to imagine how these phenomena could be explored in online experiments with large numbers of participants.

Confidence Malleability

Identification accuracy is only one component of the accuracy-confidence relationship that has received so much attention in recent scholarship (e.g., Wixted & Wells, 2017) and expert testimony (e.g., *People of the State of New York v. Boone*, 2019). Despite the legal status afforded to witness confidence as a reliable indicator by which to evaluate eyewitness testimony (Manson v. Braithwaite, 1977), witnesses' confidence statements are malleable (Stebly et al., 2014). When lineup administrators know which lineup member is the suspect, they may provide feedback to witnesses about whether they have identified the suspect. This feedback can alter witnesses' confidence in the accuracy of their identification. In addition to anecdotal evidence about police officers erupting in applause when a witness identifies a suspect (Wells, 2020), there is empirical evidence that non-blind administrators react to witness identifications by sending cues to witnesses about whether they made a "correct" identification (i.e., an identification of the suspect). For example, nonblind administrators were more likely to smile when the witness identified the suspect (Charman & Quiroz, 2016).

Confirming feedback reliably and significantly inflates witnesses' confidence in their identification accuracy (Stebly et al., 2014). Witnesses who receive confirming feedback after identifying a lineup member ("Good, you identified the suspect.") were more confident in their identification accuracy than were witnesses who received disconfirming feedback ("Actually, the suspect was ____") or no feedback (Wells & Bradfield, 1998). Repeated feedback is cumulative, with a double dose of feedback inflating witness confidence more than a single dose (Smalarz & Wells, 2020). Moreover, feedback attenuated the confidence-accuracy relationship because its effects were strongest among inaccurate eyewitnesses (Bradfield et al., 2002).

Confidence malleability is not limited to lab settings (where witnesses might presume an experimenter knows the ground truth about the guilt of a suspect); letting witnesses know whether they identified either a suspect or a filler affected confidence statements in real cases (Wright & Skagerberg, 2007). Moreover, confidence inflation can occur without administrators explicitly telling witnesses that they identified the suspect. Merely telling the witnesses that they were a good witness (Dysart et al., 2012; cf. Experiment 1; Smalarz & Wells, 2020) or providing them information that allows them to infer—without outright telling them—that they identified the suspect (Experiment 2; Smalarz & Wells, 2020) might be sufficient to inflate confidence.

Although the recommendation to use administrators who are blind to which lineup member is the suspect is—in part—to eliminate the ability of administrators to provide feedback to witnesses about their identification choices (Kovera & Evelo, 2017, 2020; Wells, Kovera et al., 2020), administrators are not the only source of feedback that has the potential to influence witness confidence statements. After hearing that a co-witness made a positive identification (i.e., had identified a lineup member but not which one) with high confidence, witnesses who also made a positive identification did so with higher confidence than did those who heard that a co-witness had made a positive identification with low confidence (Levett, 2013). Learning that a co-witness identified the same lineup member, either through notification via a computer screen (Semmler et al., 2004) or directly from the co-witness (Luus & Wells, 1994; Skagerberg, 2007), inflated witness confidence irrespective of the suspect's guilt (Semmler et al., 2004). Learning other evidence against an identified suspect, including that the suspect was identified by a jailhouse informant or that the suspect confessed to the crime, similarly increased witnesses' confidence in the accuracy of their identifications (Erickson et al., 2016; Mote et al., 2018).

Even when a contemporaneous, uncontaminated record of witness confidence is obtained, confirming feedback, misinformation about the level of confidence the witness previously expressed, and even the mere passage of time can inflate witnesses' confidence reports that are provided at a later time (Greenspan & Loftus, 2020). When witnesses are allowed to interact with others, they gain information that affects the reports of their confidence, attenuating the confidence-accuracy relationship (Bradfield et al., 2002). Therefore, in investigating questions about confidence malleability, it is important to be able to model the types of situations in which the collection of witness self-reported confidence occurs. In the real world of eyewitness behavior, these situations often involve social interactions between the witness and others.

Current Trends in Eyewitness Researchers, Theories, and Analyses

Despite the documented importance of the psychosocial context for understanding eyewitness reliability, recent trends in eyewitness research have produced conditions that reduce the likelihood that researchers will ask research questions or design studies that explore the effects of social context on witnesses. To document recent trends in the disciplinary backgrounds of scholars conducting and the theoretical frameworks and analyses used in eyewitness research, we reviewed eyewitness studies from 1980 through mid-2020 archived in the PsycINFO database. We included any empirical, peer-reviewed journal article identified through a search for the terms "eyewitness AND identification." The search returned 594 articles. We were able to locate 570 of these articles, containing 859 individual studies. We excluded any study that did not have an operationalization of eyewitness decision making as the dependent variable (e.g., jury decision-making studies about eyewitness evidence, $n = 159$) or was not a primary source (e.g., multi-study archival analyses or reviews,

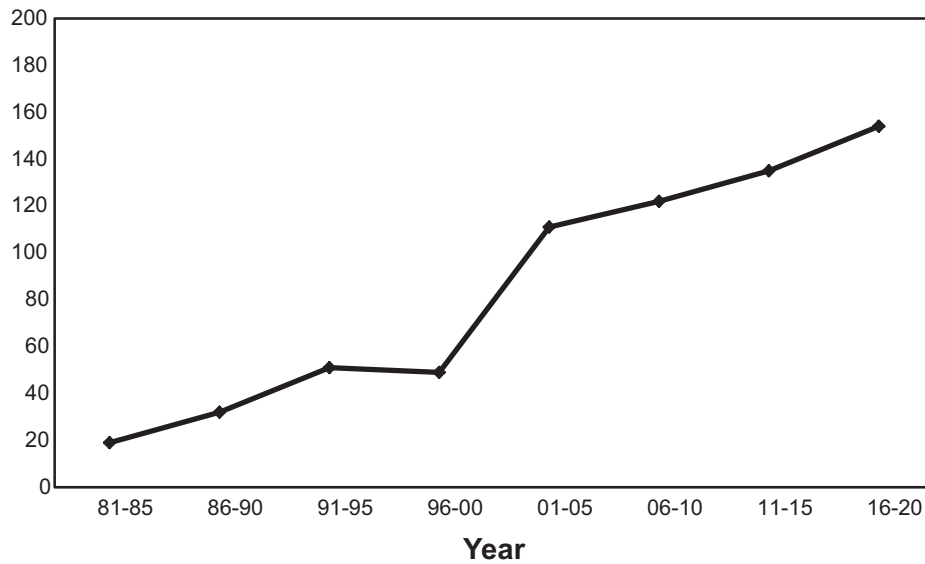


Figure 1. Number of eyewitness studies published in PsychINFO over time. Years represent year of publication of the manuscript, with each marker representing a 5-year period.

$n = 27$). The remaining population included 673 studies published from 1982 to 2020. Figure 1 shows the number of studies and the increase in eyewitness research over time. Two research assistants coded each study for whether the article measured eyewitness decision outcomes ($\kappa = .77$), the first author's background or area of training ($\kappa = .66$), if the results included SDT analysis ($\kappa = .78$), and if the study was conducted solely online ($\kappa = .74$). Disagreements were resolved by a third coder.

First, we examined whether there were any changes over time in the disciplinary backgrounds of scholars contributing to the study of eyewitness reliability. Although research in psychology and law pulls from many different psychological sub-disciplines, the topic of eyewitness investigation has generally been pursued by cognitive and social psychologists. We assessed the relative influence of these fields in the literature over time by coding the sub-disciplinary background of the first author of each of the identified papers. The proportion of studies conducted by those from a cognitive background has increased over time (see Figure 2).

The Rise of Signal Detection Theory

With the increased involvement of cognitively trained scholars in eyewitness research, it would not be surprising if the popularity of SDT frameworks also increased. The application of SDT has long been a part of eyewitness and facial recognition research to some degree. However, the earliest eyewitness studies tended only to calculate discriminability measures (e.g., Brigham et al., 1986; Shapiro & Penrod, 1986; Yarmey, 1986). Not until the 2000s did researchers use measurements of discriminability and response bias as primary tools to answer applied research questions about the effectiveness of different lineup procedures. Some argued that the perceived success of procedural reform may be the result of only changes in response bias and that psycho-legal researchers were too

focused on reducing false alarms at the cost of hits (Clark, 2012; Gronlund et al., 2015).

Cognitive psychologists also argued for the benefit of receiver operating characteristic curves (ROC; Wixted & Mickes, 2012), asserting that comparing the size of ROC curves produced by different lineup procedures was essential for the proper evaluation of potential reforms (Wixted & Mickes, 2012; Wixted et al., 2014). The National Research Council (NRC; 2014) agreed in part, stressing the importance of producing changes in discriminability rather than response bias and recommending the use of multiple statistical tools, including ROC procedures, to evaluate witness performance under different conditions. However, the application of ROC analyses to lineup performance has not been without controversy. Data collected from lineups violate the assumptions of ROC analyses. ROC analyses require binary (yes/no) or preferably continuous outcomes (Stanislaw & Todorov, 1999), but lineups can produce multiple categorical outcomes, including identifications of the suspect, identifications of fillers, lineup rejections, and "I'm not sure" responses. To avoid this violation, researchers using ROC analyses have collapsed all decisions other than an identification of the suspect into a single category, losing probative data about witness reliability (Wells et al., 2015). As a result, some have argued that ROC curves applied to lineups may not even measure underlying discriminability (e.g., Lampinen, 2016; Wells et al., 2015). However, ROC proponents have pushed back, arguing that these comments are either incorrect or irrelevant and that ROC analysis is the only proper way to analyze lineup data (Rotello & Chen, 2016; Wixted et al., 2017).

With these strong exhortations to rely on SDT and use ROC analysis for evaluating eyewitness data, it would not be surprising if the use of these theories and analyses increased over time. To test this hypothesis, our coders noted whether the authors used signal detection analysis, including measurements of discriminability, criterion, sensitivity, specificity, or area

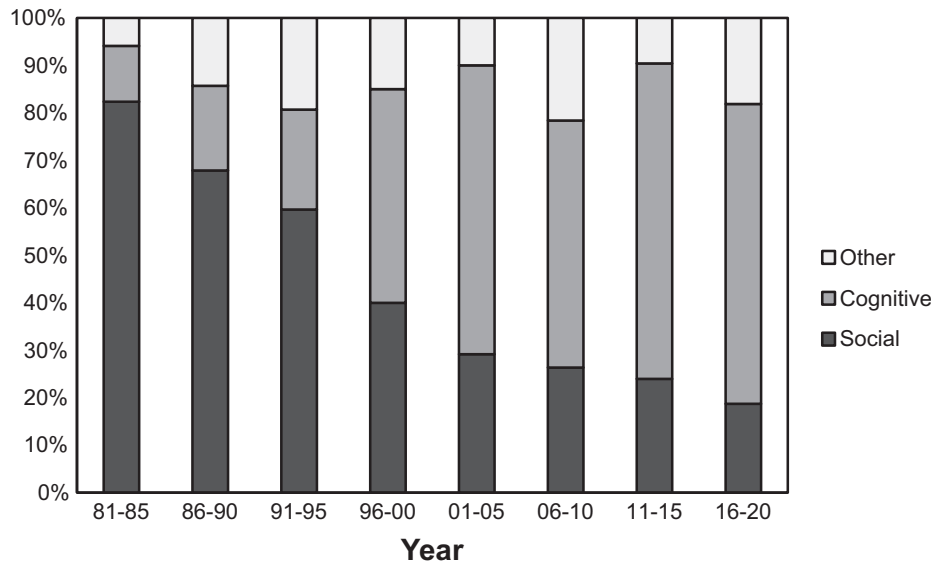


Figure 2. Percent of eyewitness studies in PsychINFO by first author background and over time. Years represent year of publication of the manuscript, with each column representing a 5-year period. First authors with a background in both social and cognitive psychology were counted toward both categories.

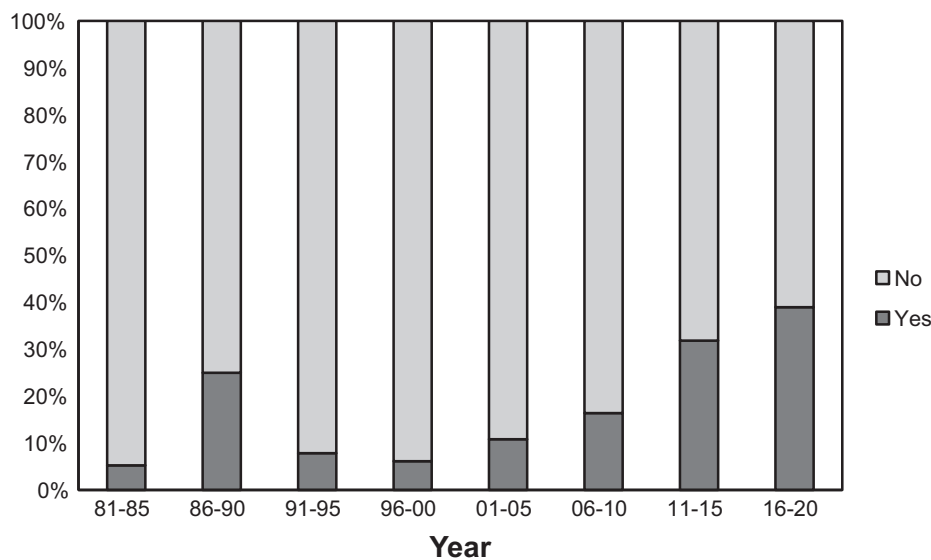


Figure 3. Percent of eyewitness studies in PsychINFO using signal-detection analyses over time. Years represent year of publication of the manuscript, with each column representing a 5-year period. Articles were counted as using signal-detection theory (“Yes”) if they used any of multiple analyses: sensitivity, specificity, criterion (c), discriminability (d'), or receiver operating characteristic (ROC). Articles were not counted toward using signal-detection theory (“No”) if they simply mentioned signal-detection theory, used signal-detection terms (e.g., hit, false-alarm) or for analyzing the confidence-accuracy relationship.

under receiver operating characteristic curves. Studies were only counted if these calculations were presented as results and not merely mentioned in theoretical discussions. Nor did we count, for the purpose of this analysis, simply labeling decisions with signal detection terminologies, such as hits or false alarms. Indeed, these analyses have become more popular over the last two decades (see Figure 3).

Increases in SDT Analyses Promote Changes in Research Methods and Questions

The application of ROC analyses to eyewitness research has changed the methods that eyewitness researchers use and the questions they ask. Methodologically, SDT raises a number

of power and precision concerns. SDT analysis can require a large number of data points to estimate the parameters correctly. This requirement is especially true for ROC analyses (Obuchowski, 2000), which estimates not one measure of discriminability, but discriminability at multiple different criteria (operationalized by using multiple false alarm rates) and—if comparing different lineup procedures—on multiple different curves. Furthermore, assuming equal precision, power in a ROC analysis is determined by the area under the curve or difference between areas under the curve (Obuchowski, 2000), both of which tend to be small in lineup studies.

Thus, a need for larger sample sizes accompanies the increased use of signal detection analyses in eyewitness research. In the articles we coded, sample sizes increased over

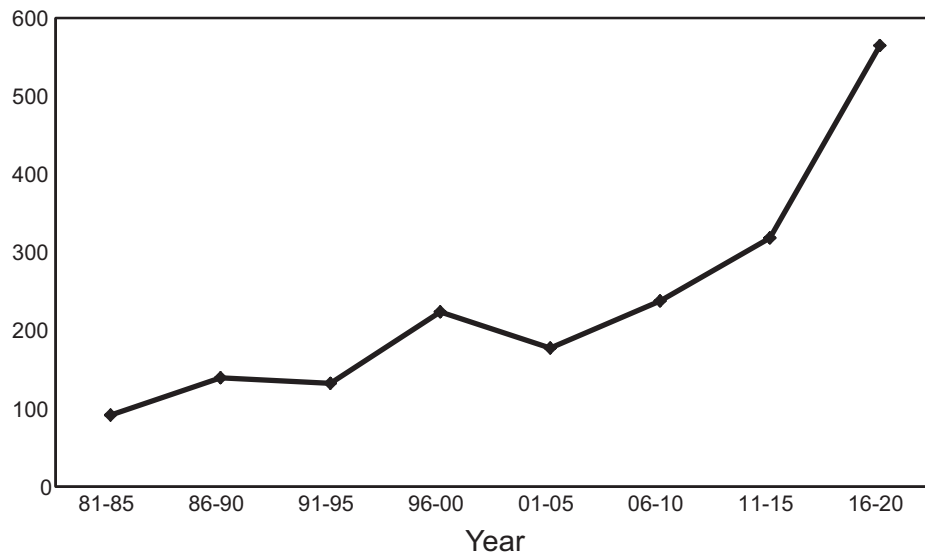


Figure 4. Mean sample size in eyewitness studies over time. Years represent year of publication of the manuscript, with each marker representing a 5-year period.

time (see Figure 4). The maximum sample size in the 1980s was 320 (Experiment 1; Cutler et al., 1986) but was 8529 in the 2010s (Colloff et al., 2016). During this past decade, methodologists warned about the sins of both *p*-hacking (Nelson et al., 2018; Simmons et al., 2011) and underpowered studies (Bakker et al., 2016), and their contribution to the replicability crisis (Lindsay, 2020). As a result, sample sizes are likely increasing across all types of psychological research. However, at least in this population of studies, there were, on average, larger sample sizes in studies using SDT analyses ($M = 627.15$, $SD = 1118.35$) than in those that did not ($M = 206.81$, $SD = 210.78$), which suggests that the increases in sample size here were related, at least in part, to the increased use of SDT analyses.

With these increases in the number of participants per eyewitness study, the field has seen a concomitant increase in the proportion of studies that collect data online from participants reached through crowdsourcing sites (e.g., Amazon's Mechanical Turk, CrowdFlower, Prolific Academic) relative to those with data collected in the lab. Figure 5 shows the rise in popularity of conducting eyewitness studies completely online in the last decade. The increase in online data collection in the eyewitness realm was disproportionately present in research using an SDT framework; authors who used an SDT framework were more likely to conduct studies only online (28.3%) than were authors who did not (4.8%).

The Unintended Consequences of Changing Analyses and Methods

Our concerns about the increasing number of participants required for ROC analyses and the concomitant boom in online data collection methods do not arise from questions about the quality of the data collected through these crowdsourcing platforms. Instead, our concerns center on the inevitable and unintended consequences that these methodological changes will have on the types of research questions that eyewitness schol-

ars pose. The ease of online data collection is bound to pressure those who wish to keep pace with others' scholarly productivity to eschew research questions that require data collection in the laboratory, often one participant at a time. As studies move from the lab to online, researchers will test questions about the variables that affect eyewitness accuracy and witness confidence necessarily using paradigms that are devoid of the social context in which real witnesses' memories are tested; studies that manipulate aspects of the social context to examine their effects on witness performance will become increasingly rare.

A preference for particular methods can lead to a lack of diversification in the research questions that scholars ask (Syed, 2021). If scholars are quickly churning out studies using online data collection and crowdsourced participants, then other scholars are likely to follow suit, which leads to what Anderson and colleagues (2019) referred to as the MTurkification of social psychology. They argued that MTurkification could lead scholars to abandon lines of inquiry that are not well-suited for online platforms, resulting in the relative extinction of whole research areas (e.g., group dynamics in social psychology). Even before the advent of crowdsourcing participants, some psychologists were bemoaning the fact that the study of psychology had become the "science of self-reports and finger movements" rather than the study of behavior (Baumeister et al., 2007, p. 396).

The same fate may be facing the study of eyewitness behavior. With pressures to publish mounting for graduate students who wish to find jobs in academia, junior faculty who wish to earn tenure, and mid-career faculty who seek promotion, the draw of rapid data collection can be intense. Couple this pressure to publish with the increased emphasis on ROC analyses as the best way to analyze eyewitness data, and online data collection becomes even more attractive. Yet, there are important areas of inquiry into eyewitness behavior that are difficult, if not impossible, to investigate on online platforms. In real cases, eyewitnesses' identification decisions and their reports of confidence in the accuracy of those decisions are not typi-

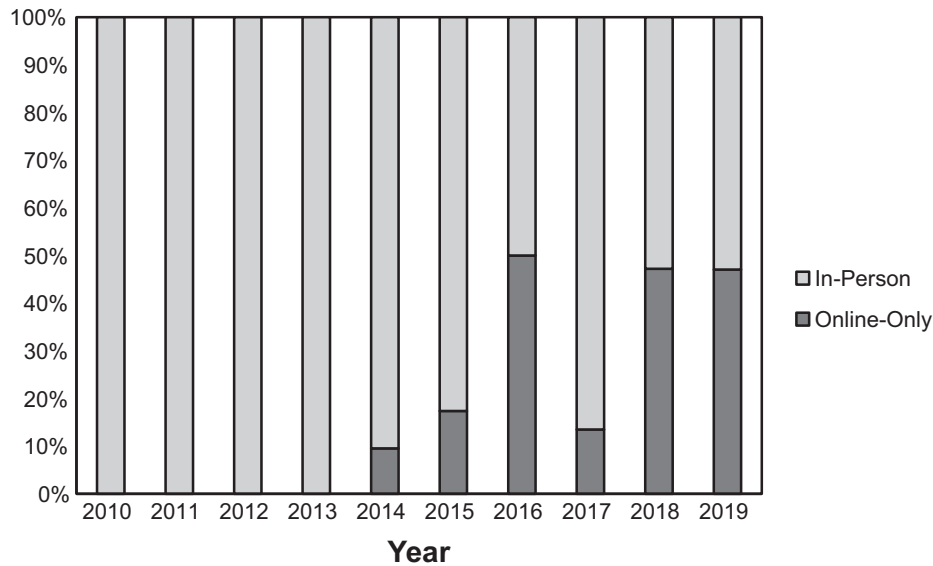


Figure 5. Percent of eyewitness studies in PsycINFO conducted entirely online over time. Time frame is for the last decade. Year represents the date of publication of the manuscript. Online only studies were completed entirely online, but in-person studies needed only to have some portion of the study in-person.

cally made online, without interaction with others. Indeed, the social context in which these decisions are made is important.

Future Directions for Studying Eyewitness Identification in Its Social Context

Given that identification accuracy and witness confidence are at the very heart of what eyewitness scholars study and that social context clearly plays a determinative role in the values these variables obtain, it is somewhat surprising that the role of social context in eyewitness behavior remains relatively understudied (Wells, 2020). After all, “every thought, feeling, desire, and behavior is embedded in a situation” (Rauthmann & Sherman, 2020, p. 473). If there remains continued emphasis on SDT models and ROC analysis, the social context effects on eyewitness behavior will likely remain understudied, especially as the number of participants required to test these models encourage the use of online data collection methods—methods that are, by their very nature, devoid of social interactions or context. Indeed, there are a number of important areas of research that have been relatively unexplored but will require labor-intensive data collection methods to investigate.

Administrator Influence

Although there is a small and growing literature on the ways in which non-blind administrators influence witness behavior (Kovera & Evelo, 2017, 2020), there is much more to learn about administrator influence more generally. For example, can administrators who are blind to which lineup member is the suspect still influence choosing? Certainly, the answer is yes. Maybe officers who know that there is not much evidence against a suspect might believe witnesses would need more encouragement to make an identification, especially as those same recommendations for best practice warn against doing repeated identification procedures with the same witness and

suspect. After all, lineup administrators’ beliefs about the difficulty of the identification task and/or the quality of a witness’s memory might influence witnesses’ identification decisions. When the same administrators—blind to which lineup member was the suspect—conducted the same identification procedure with two witnesses (the first of which was a confederate), they exerted more influence on the second witness when the confederate made a low confidence identification (Douglass et al., 2005).

Thus, even if lineup administrators are blind to which lineup member is the suspect, the interaction between the administrator and the witness may still prove problematic. For example, it would not be surprising if lineup administrators are certain of suspects’ guilt, irrespective of their actual guilt, because research on interrogations and false confessions makes clear that interrogators are guilt presumptive (Lidén et al., 2018; Meissner & Kassin, 2002). What if these guilt presumptive administrators are constrained from repeating an identification with a witness? Might they encourage the witness to make an identification from the lineup even if they have no information to convey to the witness about which lineup member to pick? In this situation, we may see administrators influence criterion or choosing. If this pressure to choose is paired with conditions that bias the witness toward choosing the suspect, like low similarity fillers, we may see an increase in mistaken identifications. Or perhaps administrators, who learn that only high confidence identifications are deemed useful by prosecutors and the courts, will encourage witnesses to inflate their self-reported confidence. If so, then it may be best to remove the social interaction from the identification procedure altogether by using computerized methods of lineup administration.

In addition to these influence attempts, which may well occur outside the awareness of the lineup administrator, there is a need to explore situations that give rise to administrators’ intentional influence attempts as well as the normative factors

that influence witnesses' willingness to comply with the administrators' wishes. There are many instances of intentional pressure placed on witnesses among the exoneration cases handled by the various branches of the Innocence Project. In most of these cases, there is no evidence other than witnesses' reports, often after the once-convicted defendant had been exonerated, that they had been pressured by police. However, there are increasing calls for the video recording of lineup administrations (Modjadidi & Kovera, 2018; Wells, Kovera et al., 2020), and some jurisdictions are now routinely doing so. In real cases that we have reviewed, there have been video records of investigators using the Reid technique (Inbau, Reid, Buckley, & Jayne, 2015)—an interrogation technique typically used to elicit confessions from suspects (Kassin et al., 2010)—to elicit identifications from reluctant witnesses.

In sum, psycho-legal researchers have only scratched the surface of the potential expectations that administrators may have about an identification procedure and its outcome. Administrators may have beliefs about the difficulty of the identification task, both because of the task itself and the conditions under which witnesses encoded their memories of the culprit, and the level of cooperation they expect to receive from a particular witness. With some of these expectations, administrators may opt to apply pressure to witnesses, and the witnesses, in turn, may make identifications that are in no way based on underlying memory. For example, if administrators believe that cross-race identifications are more difficult, they may place pressure on witnesses who differ in race or ethnicity from the suspect. It may also be that witnesses are more receptive of influence from lineup administrators with whom they share a racial, ethnic, or gender identity. There is also some evidence that the mere presence of administrators, even those who does not know which lineup member is the suspect, impairs the ability of witnesses to make cross-race identifications (Rothweiler et al., 2020). We will not be able to explain the identifications in these cases merely with a focus on discriminability or criterion.

Witness Motivations

We know very little about whether witness motivations influence witness behavior. For example, there is surprisingly little internally valid empirical evidence to allow an evaluation of whether victims of crimes make different identification decisions than do bystander witnesses, let alone whether the identification decisions of people who witness crimes in an experimental setting differ from the decisions made by people who witness real crimes. The jury decision making literature, in contrast, is replete with concerns about whether findings from video simulations generalize to juries deciding real cases, arguing that the consequentiality of the decisions varies quite a bit (for a review, see Bornstein, 2017). Moreover, both attorneys (Mecklenburg et al., 2008) and scholars (Yuille, 1993; Yuille & Cutshall, 1986; Yuille & Tollestrup, 1992) have questioned whether the findings that scholars have generated in the lab generalize to the reliability of the memories of real witnesses to real crimes. The consequentiality of the event for victims

and bystander witnesses may very well alter the motivations with which they approach an identification task. The differential effects of consequentiality on the motivations of real versus experimental witnesses may be even more stark.

When researchers have tackled the issue of whether victims and bystander identification decisions differ, they have rarely isolated variation in role from other variables. Analyses of police records sometimes show that crime victims identify more suspects than do bystander witnesses (Tollestrup et al., 1994), bystanders identify more suspects than victims (Memon et al., 2011), and sometimes show no differences (Behrman & Davey, 2001; Valentine et al., 2003). It is difficult to draw any conclusions from these archival studies because it is unclear what proportion of the identifications of suspects were correct or mistaken. Moreover, there are issues of multicollinearity in the predictors used in archival studies, with victim/bystander status systematically varying with other variables that affect the quality of encoding, like distance and presence of a weapon (Horry et al., 2014). These problems of confounding victim/bystander status with encoding variables are not unique to archival studies, with both field studies (Carlucci et al., 2011; Odinet et al., 2009) and lab studies (Hosch & Cooper, 1982; Hosch et al., 1984; Kassin, 1984) seemingly confounding witness type with other variables.

We know of only one published paper that reports the results of studies that cleanly manipulated whether the witnesses believed that they were making an identification in a real investigation or a laboratory experiment; witnesses were more likely to make a positive identification from a showup when it appeared to be part of a real investigation rather than part of a lab experiment (Experiments 1 and 2; Eisen, Smith et al., 2017). Thus, it seems likely that there are motivational differences between different types of witnesses (i.e., victim vs. bystander, real vs. laboratory) that will certainly influence their criteria for choosing but may also interact in unknown ways with other variables to affect identification decisions.

There may also be motivational differences, conscious or unconscious, that affect witness identification accuracy for members of their racial or ethnic ingroups versus outgroups. For example, the sociocultural context affects accuracy in recognizing own and other race faces. In one study, priming different cultural identities moderated the size of the own-race bias (ORB) effect, calling attention to the importance of the sociocultural context in which witnesses make identifications (Marsh, 2021). In addition, there are cultural differences in the manifestation of the ORB effect, with a significant ORB manifesting for members of cultures that essentialize (i.e., believe that racial groups are homogenous) their outgroups more than their ingroups (White Americans) but not for members of cultures that essentialize their ingroups more than their outgroups (Asian-Americans; Leffers & Coley, 2021). It is also possible that in some circumstances, witnesses may be less likely to make a positive identification of a suspect with whom they share racial group membership, much like there is an ingroup leniency effect in jury decision making (Mitchell et al., 2005). It is also possible that phenotypic bias—resulting from greater activation of racial stereotypes about criminality

by facial features that more closely resemble the African facial phenotype (Blair et al., 2002, 2004; Eberhardt et al., 2004)—could put suspects with more stereotypically African features at risk of misidentification. In sum, there is still much to learn about the role that race may play in witness behavior.

The Suspect's Behaviors

Surveys of police practice suggest that many jurisdictions use identification procedures that involve presenting a live suspect alone (a showup) or in a corporeal lineup (in which lineup members are physically present rather than depicted by showups; *Police Executive Research Forum*, 2013). These suspects and fillers are not static, passive, two-dimensional stimuli; they are living, breathing beings who behave. Suspects may be nervous, irrespective of whether they are innocent or guilty, because if they are identified by the witness, they are likely going to be indicted. Fillers may be nonchalant if they have been paid for their participation or given an hour out of their cell in exchange for their participation. Fillers could also be complicit in influencing the outcome of the lineup, either intentionally or unintentionally, in the case of fillers who are police officers who know which lineup member is the suspect.

Suspects and fillers in live procedures may emit behavioral cues that might influence witnesses' decisions. Both police officers and laypeople believe that liars emit behavioral cues, like gaze aversion and movement, despite the fact that these cues are not diagnostic of deception (Bogaard et al., 2016; Hartwig & Bond, 2011; Vrij et al., 2006). Similarly, witnesses might believe that behavioral cues suggesting that the suspect is nervous are diagnostic of suspect guilt. These beliefs might lead witnesses to make an identification, even though one could readily imagine an innocent suspect might be nervous about being mistakenly identified. And although identifications from lineups are more reliable than are identifications from showups (Colloff & Wixted, 2020; Gronlund et al., 2012; Smith et al., 2017), these studies have been conducted with photographs, not live people. What if suspects exhibit nervousness that is now more easily detected because the surrounding fillers are not nervous?

Might race play a role in the behaviors that suspects display in identification procedures? Najdowski (2011) has argued that innocent Black suspects who are being interrogated may experience stereotype threat—that is, the fear that one will be judged poorly based on a negative stereotype that is associated with one's ingroup (Steele, 2010; Steele & Aronson, 1995). The experience of stereotype threat is associated with anxiety, physiological arousal, and decreased cognitive load capabilities (Steele, 2010). Although much of the study of stereotype threat has been conducted within educational rather than criminal contexts, Black men and women are aware that people associate criminality with their racial group (Sigelman & Tuch, 1997). Moreover, Black men express concern that police will judge them negatively in a criminal context because of their race, and Black men are more likely than White men to report that in an encounter with the police they would be anxious and exhibit behaviors that police find suspicious (Najdowski et al.,

2015). The exhibition of these suspicious behaviors might also place Black men at higher risk than White men of being positively identified in a showup or corporeal lineup.

Conclusion

As most oenophiles will tell you, the shape of a glass changes the flavor of the wine it contains (Newton, 2015). Similarly, psychologists' preferred theories and methods can constrain and change the types of research questions that they ask (Syed, 2021). With the increasing popularity of SDT frameworks, ROC analyses, and online data collection methods, many recent studies fail to capture the social context in which eyewitness identifications occur. There are many impediments to doing studies that manipulate or measure behavior that have been discussed more exhaustively elsewhere (Anderson et al., 2019; Baumeister et al., 2007). Methods that either manipulate or measure real behavior are costly and time-intensive, often allowing for the collection of only one data point per experimental session. Given the importance of appropriately powering our statistical tests, highlighted by the ongoing replication crisis in psychology (Maxwell et al., 2015), these studies become Herculean efforts. Because evaluative criteria in academia often give heavy weight to publication quantity, as a field we will need to recognize the additional effort and resources required for research that investigates eyewitness behavior in its social context when evaluating candidates for academic jobs or faculty for tenure, promotion, grants, and awards. Without this recognition, the study of eyewitness behavior in its social context will wither, much like the study of group dynamics and behavior has in the field of social psychology (Anderson et al., 2019).

We also caution against presenting fact finders with the conclusion that identifications made with high confidence will be accurate when both confidence statements and identifications were collected under a specific set of pristine conditions (Wixted & Wells, 2017). We are not contesting that there is a strong relationship between accuracy and confidence in situations that are devoid of social context. However, our review makes clear that there are social influences on witnesses' identification accuracy and self-reported confidence, like co-witness information, that are not eliminated by pristine conditions. Until studies of eyewitness behavior test whether this relationship holds across a variety of settings, demonstrating the generalizability of the findings to contexts in which eyewitnesses are making their judgments, we believe it is premature to opine that highly confident witnesses are necessarily accurate.

These observations are not intended, in any way, to detract from the important contributions that cognitive psychologists, their theories, methods, and analyses have made to advancing our understanding of eyewitness memory, including their identification decisions and confidence in those decisions. The introduction of SDT frameworks and analyses has made a substantial contribution to recent advances in eyewitness research. SDT has sharpened the analysis of the abilities of eyewitnesses to accurately identify perpetrators of crime based on their memories, increased the use of theoretical and predictive models, standardized important terminology across disciplines,

and brought attention to the predictive utility of witness confidence.

Nor are these observations meant to discourage researchers from studying eyewitness memory, online or otherwise. Understanding how eyewitness memory works is important, but the use of SDT and ROC analysis is not the *sine qua non* of eyewitness research. Science thrives through the use of multiple methods and a diversity of viewpoints and theories. However, theories based on an SDT framework (e.g., Wixted & Mickes, 2014) are likely impractical for understanding eyewitness behavior. A practical theory is one that starts with a problem, refines itself in an iterative process involving practitioners and scientists, and engages with data drawn from real-world observation of the phenomenon it wishes to describe (Berkman & Wilson, 2021; Lewin, 1943). Observations of police practices in real cases suggest that understanding memory will only sometimes be relevant for understanding eyewitness behavior. A practical theory of eyewitness behavior will allow researchers to specify the conditions under which social context is an important predictor of accuracy. As Neisser (1978, p. 4) wrote decades ago, “the naturalistic study of memory is an idea whose time has come.”

Author Contributions

M.B.K. conceived of the argument presented in the paper. A. J.E. oversaw the coding of the articles and the preparation of the first draft of the section presenting the data on trends in eyewitness scholarship. M.B.K. wrote the first draft of the remaining sections and made changes to the manuscript in light of reviewer comments. Both authors contributed to the final preparation of the manuscript and agreed upon the final version.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by grants from the National Science Foundation (NSF) to the first author under Grants SES-9986240, SES-0922314, and SES-2043334 and the first and second authors under Grant Number SES-1728938. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. We are grateful to Eliana Aronson, Miriam Lieber, and Dasi Richardson for their assistance with the coding of the data.

References

- Alceste, F., Jones, K. A., & Kassin, S. M. (2020). Facts only the perpetrator could have known? A study of contamination in mock crime interrogations. *Law and Human Behavior*, 44, 128–142. <https://doi.org/10.1037/lhb0000367>.
- Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., & Rokkum, J. N. (2019). The MTurkification of social and

- personality psychology. *Personality and Social Psychology Bulletin*, 45, 842–850. <https://doi.org/10.1177/0146167218798821>.
- Bakker, M., Hartgerink, C. H. J., Wicherts, J. M., & van der Maas, H. L. J. (2016). Researchers' intuitions about power in psychological research. *Psychological Science*, 27, 1069–1077. <https://doi.org/10.1177/0956797616647519>.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?. *Perspectives on Psychological Science*, 2, 396–403. <https://doi.org/10.1111/j.1745-6916.2007.00051.x>.
- Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, 25, 475–491. <https://doi.org/10.1023/A:1012840831846>.
- Berkman, E. T., & Wilson, S. M. (2021). So useful as a good theory? The practicality crisis in (social) psychological theory. *Perspectives on Psychological Science*. Advance online publication. <https://doi.org/10.1177/1745691620969650>.
- Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology*, 87, 763–778. <https://doi.org/10.1037/0022-3514.87.6.763>.
- Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology*, 83, 5–25. <https://doi.org/10.1037/0022-3514.83.1.5>.
- Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PLoS ONE*, 11, e0156615. <https://doi.org/10.1371/journal.pone.0156615>.
- Bornstein, B. H. (2017). Jury simulation research: Pros, cons, trends, and alternatives. In M. B. Kovera (Ed.), *The psychology of juries* (pp. 207–226). American Psychological Association. <https://doi.org/10.1037/0000026-010>.
- Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology*, 87, 112–120. <https://doi.org/10.1037/0021-9010.87.1.112>.
- Brewin, C. R., Andrews, B., & Mickes, L. (2020). Regaining consensus on the reliability of memory. *Current Directions in Psychological Science*, 29, 121–125. <https://doi.org/10.1177/0963721>.
- Brigham, J. C., Maass, A., Martinez, D., & Whittenberger, G. (1986). The effect of arousal on facial recognition. *Basic and Applied Psychology*, 4, 279–293. https://doi.org/10.1207/s15324834basp0403_6.
- Carlson, C. A., Dias, J. L., Weatherford, D. R., & Carlson, M. A. (2017). An investigation of the weapon focus effect and the confidence-accuracy relationship for eyewitness identification. *Journal of Applied Research in Memory and Cognition*, 6, 82–92. <https://doi.org/10.1016/j.jarmac.2016.04.001>.
- Carlucci, M. E., Kieckhafer, J. M., Schwartz, S. L., Villalba, D. K., & Wright, D. B. (2011). The South Beach study: Bystanders' memories are more malleable. *Applied Cognitive Psychology*, 25, 562–566. <https://doi.org/10.1002/acp.v25.410.1002/acp.1720>.
- Charman, S. D., & Quiroz, V. (2016). Blind sequential lineup administration reduces both false identifications and confidence in those identifications. *Law and Human Behavior*, 40, 477–487. <https://doi.org/10.1037/lhb0000197>.

- Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science*, 7, 238–259. <https://doi.org/10.1177/1745691612439584>.
- Clark, S. E., Brower, G. L., Rosenthal, R., Hicks, J. M., & Moreland, M. B. (2013). Lineup administrator influences on eyewitness identification and eyewitness confidence. *Journal of Applied Research in Memory and Cognition*, 2, 158–165. <https://doi.org/10.1016/j.jarmac.2013.06.003>.
- Clark, S. E., Marshall, T. E., & Rosenthal, R. (2009). Lineup administrator influences on eyewitness identification decisions. *Journal of Experimental Psychology: Applied*, 15, 63–75. <https://doi.org/10.1037/a0015185>
- Colloff, M. F., Wade, K. A., & Strange, D. (2016). Unfair lineups make witnesses more likely to confuse innocent and guilty suspects. *Psychological Science*, 27, 1227–1239. <https://doi.org/10.1177/0956797616655789>.
- 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, 124–143. <https://doi.org/10.1037/xap0000218>.
- Cutler, B. L., Penrod, S. D., O'Rourke, T. E., & Martens, T. K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour*, 1, 113–134.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51, 629–636. <https://doi.org/10.1037/h0046408>.
- Douglass, A. B., Lucas, C. A., & Brewer, N. (2020). Cowitness identification speed affects choices from target-absent photo-spreads. *Law and Human Behavior*, 44, 474–484. <https://doi.org/10.1037/lhb0000420>.
- Douglass, A. B., Smith, C., & Fraser-Thill, R. (2005). A problem with double-blind photospread procedures: Photospread administrators use one eyewitness's confidence to influence the identification of another eyewitness. *Law and Human Behavior*, 29, 543–562. <https://doi.org/10.1007/s10979-005-6830-9>.
- Dysart, J. E., Lawson, V. Z., & Rainey, A. (2012). Blind lineup administration as a prophylactic against the post-identification feedback effect. *Law and Human Behavior*, 36, 312–319. <https://doi.org/10.1037/h0093921>.
- Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, 87, 876–893. <https://doi.org/10.1037/0022-3514.87.6.876>.
- Eisen, M. L., Gabbert, G., Ying, R., & Williams, J. (2017). “I think he had a tattoo on his neck”: How co-witness discussions about a perpetrator's description can affect eyewitness identification decisions. *Journal of Applied Research in Memory and Cognition*, 6, 274–282. <https://doi.org/10.1016/j.jarmac.2017.01.009>.
- Eisen, M. L., Smith, A. M., Olaguez, A. P., & Skerritt-Perta, A. S. (2017). An examination of showups conducted by law enforcement using a field-simulation paradigm. *Psychology, Public Policy, and Law*, 23, 1–22. <https://doi.org/10.1037/law0000115>.
- Erickson, W. B., Lampinen, J. M., Wooten, A., Wetmore, S., & Neuschatz, J. (2016). When snitches corroborate: Effects of post-identification feedback from a potentially compromised source. *Psychiatry, Psychology and Law*, 23, 148–160. <https://doi.org/10.1080/13218719.2015.1035623>.
- Evelo, A. J. (2020). *Extra-legal information transfer during eyewitness identification*. [Doctoral dissertation, City University of New York]. CUNY Academic Works. https://academicworks.cuny.edu/gc_etds/3750.
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology*, 17, 533–543. [https://doi.org/10.1002/\(ISSN\)1099-072010.1002/acp.v17:510.1002/acp.885](https://doi.org/10.1002/(ISSN)1099-072010.1002/acp.v17:510.1002/acp.885).
- Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology*, 9, 215–227. <https://doi.org/10.1348/1355325041719428>.
- Greathouse, S. M., & Kovera, M. B. (2009). Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification. *Law and Human Behavior*, 33, 70–82. <https://doi.org/10.1007/s10979-008-9136-x>.
- Greenspan, R. L., & Loftus, E. F. (2020). Eyewitness confidence malleability: Misinformation as post-identification feedback. *Law and Human Behavior*, 44, 194–208. <https://doi.org/10.1037/lhb0000369>.
- Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., & Graham, M. (2012). Shoups versus lineups: An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition*, 1, 221–228. <https://doi.org/10.1016/j.jarmac.2012.09.003>.
- Gronlund, S. D., Mickes, L., Wixted, J. T., & Clark, S. E. (2015). Conducting an eyewitness lineup: How the research got it wrong. In N. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 63, pp. 1–43). Elsevier.
- Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, 137, 643–659. <https://doi.org/10.1037/a0023589>.
- Horry, R., Halford, P., Brewer, N., Milne, R., & Bull, R. (2014). Archival analyses of eyewitness identification test outcomes: What can they tell us about eyewitness memory? *Law and Human Behavior*, 38, 94–108. <https://doi.org/10.1037/lhb0000060>.
- Hosch, H. M., & Cooper, D. S. (1982). Victimization as a determinant of eyewitness accuracy. *Journal of Applied Psychology*, 67, 649–652. <https://doi.org/10.1037/0021-9010.67.5.649>.
- Hosch, H. M., Leippe, M. R., Marchioni, P. M., & Cooper, S. (1984). Victimization, self-monitoring, and eyewitness identification. *Journal of Applied Psychology*, 69, 280–288. <https://doi.org/10.1037/0021-9010.69.2.280>.
- Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2015). *Essentials of the Reid Technique: Criminal interrogation and confessions* (2nd ed.). Jones & Bartlett Learning.
- Jack, F., Zydervelt, S., & Zajac, R. (2014). Are co-witnesses special? Comparing the influence of co-witness and interviewer misinformation on eyewitness reports. *Memory*, 22, 243–255. <https://doi.org/10.1080/09658211.2013.778291>.
- Kassin, S. M. (1984). Eyewitness identification: Victims versus bystanders. *Journal of Applied Social Psychology*, 14, 519–529. <https://doi.org/10.1111/jasp.1984.14.issue-610.1111/j.1559-1816.1984.tb02257.x>.
- Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, 34, 3–38. <https://doi.org/10.1007/s10979-009-9188-6>.

- Kovera, M. B., & Evelo, A. J. (2017). The case for double-blind lineup administration. *Psychology, Public Policy, and Law*, 23, 421–437. <https://doi.org/10.1037/law0000139>.
- Kovera, M. B., & Evelo, A. J. (2020). Improving eyewitness identification evidence through double-blind lineup administration. *Current Directions in Psychological Science*, 29, 563–568. <https://doi.org/10.1177/0963721420969366>.
- Lampinen, J. M. (2016). ROC analyses in eyewitness identification research. *Journal of Applied Research in Memory and Cognition*, 5, 21–33. <https://doi.org/10.1016/j.jarmac.2015.08.006>.
- Leffers, J. S., & Coley, J. D. (2021). Do I know you? The role of culture in racial essentialism and facial recognition memory. *Journal of Applied Research in Memory and Cognition*, 10, 5–12. <https://doi.org/10.1016/j.jarmac.2020.12.001>.
- Levett, L. M. (2013). Co-witness information influences whether a witness is likely to choose from a lineup. *Legal and Criminological Psychology*, 18, 168–180. <https://doi.org/10.1111/j.2044-8333.2011.02033.x>.
- Lewin, K. (1943). Psychology and the process of group living. *Journal of Social Psychology*, 17, 113–131. <https://doi.org/10.1080/00224545.1943.9712269>.
- Lidén, M., Gräns, M., & Juslin, P. (2018). The presumption of guilt in suspect interrogations: Apprehension as a trigger of confirmation bias and debiasing techniques. *Law and Human Behavior*, 42, 336–354. <https://doi.org/10.1037/lhb0000287>.
- Lindsay, D. S. (2020). Seven steps toward transparency and replicability in psychological science. *Canadian Psychology*, 61, 310–317. <https://doi.org/10.1037/cap0000222>.
- 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, 1047–1060. <https://doi.org/10.1002/acp.3683>.
- Luus, C. A. E., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, 79, 714–723. <https://doi.org/10.1037/0021-9010.79.5.714>.
- Macmillan, N. A., & Creelman, C. D. (2010). *Detection theory: A user's guide* (2nd ed.). Psychological Press, Taylor & Francis Group.
- Manson v. Braithwaite, 432 U.S. 98 (1977).
- Marsh, B. U. (2021). The cost of racial salience on face memory: How the cross-race effect is moderated by racial ambiguity and the race of the perceiver and the perceived. *Journal of Applied Research in Memory and Cognition*, 10, 13–23. <https://doi.org/10.1016/j.jarmac.2020.09.008>.
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist*, 70, 487–498. <https://doi.org/10.1037/a0039400>.
- Mecklenburg, S. H., Bailey, P. J., & Larson, M. R. (2008). The Illinois Field Study: A significant contribution to understanding real world eyewitness identification issues. *Law and Human Behavior*, 32, 22–27. <https://doi.org/10.1007/s10979-007-9108-6>.
- Meissner, C. A., & Kassin, S. M. (2002). “He’s guilty!”: Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26, 469–480. <https://doi.org/10.1023/A:1020278620751>.
- Memon, A., Havard, C., Clifford, B., Gabbert, F., & Watt, M. (2011). A field evaluation of the VIPER system: A new technique for eliciting eyewitness evidence. *Psychology, Crime, & Law*, 17, 711–729. <https://doi.org/10.1080/10683160903524333>.
- Mitchell, T. L., Haw, R. M., Pfeifer, J. E., & Meissner, C. A. (2005). Racial bias in mock juror decision-making: A meta-analytic review of defendant treatment. *Law and Human Behavior*, 29, 621–637. <https://doi.org/10.1007/s10979-005-8122-9>.
- Modjadidi, K., & Kovera, M. B. (2018). Viewing videotaped identification procedure increases jurors’ sensitivity to single-blind lineup administration. *Law and Human Behavior*, 42, 244–257. <https://doi.org/10.1037/lhb0000288>.
- Mote, P. M., Neuschatz, J. S., Bornstein, B. H., Wetmore, S. A., & Key, K. N. (2018). Secondary confessions as post-identification feedback: How jailhouse informant testimony can alter eyewitnesses’ identification decisions. *Journal of Police and Criminal Psychology*, 33, 375–384. <https://doi.org/10.1007/s11896-018-9274-0>.
- National Research Council (2014). *Identifying the culprit: Assessing eyewitness identification*. The National Academy Press. <https://doi.org/10.17226/18891>.
- Najdowski, C. J. (2011). Stereotype threat in criminal interrogations: Why innocent Black suspects are at risk for confessing falsely. *Psychology, Public Policy, and Law*, 17, 562–591. <https://doi.org/10.1037/a0023741>.
- Najdowski, C. J., Bottoms, B. L., & Goff, P. A. (2015). Stereotype threat and racial differences in citizens’ experiences of police encounters. *Law and Human Behavior*, 39, 463–477. <https://doi.org/10.1037/lhb0000140>.
- Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). Academic Press.
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology’s renaissance. *Annual Review of Psychology*, 69, 511–534. <https://doi.org/10.1146/annurev-psych-122216-011836>.
- Newton, J. (2015, April). *Wine snobs are right: Glass shape does affect flavor*. Chemistry World. <https://www.chemistryworld.com/news/camera-turned-wine-connoisseur/8443.article>
- Obuchowski, N. A. (2000). Sample size tables for receiver operating characteristic studies. *American Journal of Roentgenology*, 175, 603–608. <https://doi.org/10.2214/ajr.175.3.1750603>.
- Odinot, G., Wolgers, G., & van Koppen, P. J. (2009). Eyewitness memory of a supermarket robbery: A case study of accuracy and confidence after 3 months. *Law and Human Behavior*, 33, 506–514. <https://doi.org/10.1007/s10979-008-9152-x>.
- Paterson, H. M., & Kemp, R. I. (2006). Comparing methods of presenting post-event information: The power of co-witness suggestion. *Applied Cognitive Psychology*, 20, 1083–1099. <https://doi.org/10.1002/acp.1261>.
- People of the State of New York v. Boone, 2190-11 (2019).
- Pezdek, K., Abed, E., & Cormia, A. (2020). Elevated stress impairs the accuracy of eyewitness memory but not the confidence–accuracy relationship. *Journal of Experimental Psychology: Applied*, 27, 158–169. <https://doi.org/10.1037/xap0000316>.
- Police Executive Research Forum. (2013). *A national survey of eyewitness identification procedures in law enforcement agencies*. http://www.policeforum.org/assets/docs/Free_Online_Documents/Eyewitness_Identification/a%20national%20survey%20of%20eyewitness%20identification%20procedures%20in%20law%20enforcement%20agencies%202013.pdf
- Rauthmann, J. F., & Sherman, R. A. (2020). The situation of situation research: Knowns and unknowns. *Current Directions in Psycho-*

- logical Science, 29, 473–480. <https://doi.org/10.1177/0963721420925546>.
- Rotello, C. M., & Chen, T. (2016). ROC curve analyses of eyewitness identification decisions: An analysis of the recent debate. *Cognitive Research: Principles and Implications*, 1, 10. <https://doi.org/10.1186/s41235-016-0006-7>.
- Rothweiler, J. N., Goodwin, K. A., & Kukucka, J. (2020). Presence of administrators differentially impacts eyewitness discriminability for same- and other-race identifications. *Applied Cognitive Psychology*, 34, 1530–1537. <https://doi.org/10.1002/acp.3733>.
- Semmler, C., Dunn, J., Mickes, L., & Wixted, J. T. (2018). The role of estimator variables in eyewitness identification. *Journal of Experimental Psychology: Applied*, 24, 400–415. <https://doi.org/10.1037/xap0000157>.
- Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of postidentification feedback on eyewitness identification and non-identification confidence. *Journal of Applied Psychology*, 89, 334–346. <https://doi.org/10.1037/0021-9010.89.2.334>.
- Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139–156. <https://doi.org/10.1037/0033-2909.100.2.139>.
- Sigelman, L., & Tuch, S. A. (1997). Metastereotypes: Blacks' perceptions of Whites' stereotypes of Blacks. *Public Opinion Quarterly*, 61, 87–101. <https://doi.org/10.1086/297788>.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366. <https://doi.org/10.1177/0956797611417632>.
- Skagerberg, E. M. (2007). Co-witness feedback in line-ups. *Applied Cognitive Psychology*, 21, 489–497. <https://doi.org/10.1002/acp.1285>.
- Smalarz, L., & Wells, G. L. (2020). Do multiple doses of feedback have cumulative effects on eyewitness confidence? *Journal of Applied Research in Memory and Cognition*, 9, 505–518. <https://doi.org/10.1016/j.jarmac.2020.06.003>.
- 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, 127–145. <https://doi.org/10.1037/lhb0000219>.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior, Research Methods, Instruments, & Computers*, 31, 137–149. <https://doi.org/10.3758/BF03207704>.
- 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, 1–18. <https://doi.org/10.1037/law0000001>.
- Steele, C. M. (2010). *Whistling Vivaldi and other clues to how stereotypes affect us*. Norton.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>.
- Syed, M. (2021). *It's 2x2 designs all the way down: Social psychology's over-reliance on experiments restricts diversity in the field*. [Paper presentation]. Society for Personality and Social Psychology, Virtual. <https://osf.io/gc3mu/>
- Thorley, C., & Kumar, D. (2017). Eyewitness susceptibility to co-witness misinformation is influenced by co-witness confidence and own self-confidence. *Psychology, Crime & Law*, 23, 342–360. <https://doi.org/10.1080/1068316X.2016.1258471>.
- Tollestrup, P., Turtle, J., & Yuille, J. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 144–160). Cambridge University Press.
- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, 17, 969–993. [https://doi.org/10.1002/\(ISSN\)1099-0720.1002/acp.v17:810.1002/acp.939](https://doi.org/10.1002/(ISSN)1099-0720.1002/acp.v17:810.1002/acp.939).
- Vrij, A., Akehurst, L., & Knight, S. (2006). Police officers', social workers', teachers' and the general public's beliefs about deception in children, adolescents and adults. *Legal and Criminological Psychology*, 11, 297–312. <https://doi.org/10.1348/135532505X60816>.
- Wells, G. L. (2020). Psychological science on eyewitness identification and its impact on police practices and policies. *American Psychologist*, 75, 1316–1329. <https://doi.org/10.1037/amp00007>.
- Wells, G. L., & Bradfield, A. L. (1998). 'Good, you identified the suspect': Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, 83, 360–376. <https://doi.org/10.1037/0021-9010.83.3.360>.
- Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, 44, 3–36. <https://doi.org/10.1037/lhb0000359>.
- Wells, G. L., Smalarz, L., & Smith, A. M. (2015). ROC analysis of lineups does not measure underlying discriminability and has limited value. *Journal of Applied Research in Memory and Cognition*, 4, 313–317. <https://doi.org/10.1016/j.jarmac.2015.08.008>.
- Wixted, J. T., Gronlund, S. D., & Mickes, L. (2014). Policy regarding the sequential lineup is not informed by probative value but is informed by receiver operating characteristic analysis. *Current Directions in Psychological Science*, 23, 17–18. <https://doi.org/10.1177/0963721413510934>.
- Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives on Psychological Science*, 7, 275–278. <https://doi.org/10.1177/1745691612442906>.
- Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic feature-detection model of eyewitness identification. *Psychological Review*, 121, 262–276. <https://doi.org/10.1037/a0035940>.
- Wixted, J. T., Mickes, L., Wetmore, S. A., Gronlund, S. D., & Neuschatz, J. S. (2017). ROC analysis in theory and practice. *Journal of Applied Research in Memory and Cognition*, 6, 343–351. <https://doi.org/10.1016/j.jarmac.2016.12.002>.
- Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, 18, 10–65. <https://doi.org/10.1177/1529100616686966>.
- Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science*, 18, 174–178. <https://doi.org/10.1111/j.1467-8721.2009.01631.x>.

- Wright, D. B., & Skagerberg, E. M. (2007). Postidentification feedback affects real eyewitnesses. *Psychological Science*, 18, 172–178. <https://doi.org/10.1111/j.1467-9280.2007.01868.x>.
- Yarmey, A. D. (1986). Verbal, visual, and voice identification of a rape suspect under different levels of illumination. *Journal of Applied Psychology*, 71, 363–370. <https://doi.org/10.1037/0021-9010.71.3.363>.
- Yuille, J. C. (1993). We must study forensic eyewitnesses to know about them. *American Psychologist*, 48, 572–573. <https://doi.org/10.1037/0003-066X.48.5.572>.
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology*, 71, 291–301. <https://doi.org/10.1037/0021-9010.71.2.291>.
- Yuille, J. C., & Tollestrup, P. A. (1992). A model of the diverse effects of emotion on eyewitness memory. In S.-Å., Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (pp. 201–215). Erlbaum.
- Zajac, R., Dickson, J., Munn, R., & O'Neill, S. (2016). Trusst me, I know what I sshaw: The acceptance of misinformation from an apparently unreliable co-witness. *Legal and Criminological Psychology*, 21, 127–140. <https://doi.org/10.1111/lcrp.12032>.
- Zajac, R., & Henderson, N. (2009). Don't it make my brown eyes blue: Co-witness misinformation about a target's appearance can impair target-absent line-up performance. *Memory*, 17, 266–278. <https://doi.org/10.1080/09658210802623950>.
- Zimmerman, D. M., Chorn, J. A., Rhead, L. M., Evelo, A. J., & Kovera, M. B. (2017). Memory strength and lineup presentation moderate effects of administrator influence on mistaken identifications. *Journal of Experimental Psychology: Applied*, 23, 460–473. <https://doi.org/10.1037/xap0000147>.

Received February 14, 2021

Received in revised form April 5, 2021

Accepted April 11, 2021

Available online 20 August 2021