Eyewitness Evidence

Improving Its Probative Value

Gary L. Wells,[1] Amina Memon,[2] and Steven D. Penrod[3]

[1]Iowa State University; [2]University of Aberdeen, Aberdeen, Scotland; and [3]John Jay College of Criminal Justice

SUMMARY—*The criminal justice system relies heavily on eyewitnesses to determine the facts surrounding criminal events. Eyewitnesses may identify culprits, recall conversations, or remember other details. An eyewitness who has no motive to lie is a powerful form of evidence for jurors, especially if the eyewitness appears to be highly confident about his or her recollection. In the absence of definitive proof to the contrary, the eyewitness's account is generally accepted by police, prosecutors, judges, and juries.*

However, the faith the legal system places in eyewitnesses has been shaken recently by the advent of forensic DNA testing. Given the right set of circumstances, forensic DNA testing can prove that a person who was convicted of a crime is, in fact, innocent. Analyses of DNA exoneration cases since 1992 reveal that mistaken eyewitness identification was involved in the vast majority of these convictions, accounting for more convictions of innocent people than all other factors combined. We review the latest figures on these DNA exonerations and explain why these cases can only be a small fraction of the mistaken identifications that are occurring.

Decades before the advent of forensic DNA testing, psychologists were questioning the validity of eyewitness reports. Hugo Münsterberg's writings in the early part of the 20th century made a strong case for the involvement of psychological science in helping the legal system understand the vagaries of eyewitness testimony. But it was not until the mid- to late 1970s that psychologists began to conduct programmatic experiments aimed at understanding the extent of error and the variables that govern error when eyewitnesses give accounts of crimes they have witnessed. Many of the experiments conducted in the late 1970s and throughout the 1980s resulted in articles by psychologists that contained strong warnings to the legal system that eyewitness evidence was being overvalued by the justice system in the sense that its impact on triers of fact (e.g., juries) exceeded its probative (legal-proof) value. Another message of the research was that the validity of eyewitness reports depends a great deal on the procedures that are used to obtain those reports and that the legal system was not using the best procedures.

Although defense attorneys seized on this nascent research as a tool for the defense, it was largely ignored or ridiculed by prosecutors, judges, and police until the mid 1990s, when forensic DNA testing began to uncover cases of convictions of innocent persons on the basis of mistaken eyewitness accounts. Recently, a number of jurisdictions in the United States have implemented procedural reforms based on psychological research, but psychological science has yet to have its fullest possible influence on how the justice system collects and interprets eyewitness evidence.

The psychological processes leading to eyewitness error represent a confluence of memory and social-influence variables that interact in complex ways. These processes lend themselves to study using experimental methods. Psychological science is in a strong position to help the criminal justice system understand eyewitness accounts of criminal events and improve their accuracy. A subset of the variables that affect eyewitness accuracy fall into what researchers call system variables, *which are variables that the criminal justice system has control over, such as how eyewitnesses are instructed before they view a lineup and methods of interviewing eyewitnesses. We review a number of system variables and describe how psychological scientists have translated them into procedures that can improve the probative value of eyewitness accounts. We also review* estimator variables, *variables that affect eyewitness accuracy but over which the system has no control, such as cross-race versus within-race identifications.*

We describe some concerns regarding external validity and generalization that naturally arise when moving from the laboratory to the real world. These include issues of base rates, multicollinearity, selection effects, subject populations, and psychological realism. For each of these concerns, we briefly note ways in which both theory and field data help make the case for generalization.

Direct correspondence to Gary L. Wells, Psychology Department, Iowa State University, Ames, IA 50011; e-mail: glwells@iastate.edu.

INTRODUCTION

Kirk Bloodsworth had never been in trouble with the law, and yet he was convicted in March 1985 for the 1984 sexual assault and slaying of a 9-year-old girl in Maryland (*State of Maryland v. Kirk N. Bloodsworth*, 1984). Five eyewitnesses identified Bloodsworth at trial. Later that month, a judge sentenced him to death. He spent 2 years on death row before he received a new trial based on the prosecution's withholding of information about other suspects. This time he received a life sentence. Bloodsworth maintained a claim of innocence from the outset, but it was not until 1993 that he was released from prison on the basis of DNA testing that proved he was not the source of semen found in the little girl's underwear. Bloodsworth was lucky that the underwear had been preserved, because earlier (pre-DNA) tests had indicated nothing of value on the underwear. But what kind of luck is being convicted of a murder you did not commit? His mother died while he was in prison, before learning the truth that he was innocent. And despite his release from prison, some people, including one of the original prosecutors, continued to believe that Bloodsworth may have been the murderer. The eyewitness evidence just seemed too strong. Maybe Bloodsworth really was the murderer, they reasoned, and the tiny speck of semen came from someone other than the murderer—perhaps someone who had access to the little girl's dresser drawer, for instance. Bloodsworth went on with his life, confident in his own innocence but having to live with the occasional doubt raised by those who somehow remained unpersuaded. Then, in September 2003, DNA testing got a hit on the actual murderer, Kimberly Shay Ruffner. Nineteen years after Kirk Bloodsworth was sentenced to death, the proof was finally there: He had had nothing to do with the sexual assault and slaying of the young girl.

The case of Kirk Bloodsworth illustrates several problems with eyewitness evidence. First, it illustrates the fallacy of assuming that inter-witness agreement is necessarily strong evidence of accuracy. Many factors can lead to inter-witness agreement, such as interaction among the witnesses in which they share information. In general, factors that lead one eyewitness to make a particular error will lead others to make the same error. Second, the Bloodsworth case illustrates the profound level of proof required for exonerating evidence to trump eyewitness identification evidence. Even when the semen was proved not to match Bloodsworth's DNA, many people were unwilling to believe he was innocent. It was necessary to prove that someone else had committed the murder. Third, the Bloodsworth case illustrates that mistaken identification is a dual problem: Not only might an innocent person be convicted but the guilty party remains free to reoffend.

The role of scientific psychology in the problem of eyewitness evidence is a profound one. With few exceptions, the legal system has not conducted research on eyewitness evidence, has never conducted an experiment on memory, and has no scientific theory regarding how memory works. The scientific study of eyewitnesses is purely the domain of psychology. When the U.S. Department of Justice finally wrote guidelines on eyewitness evidence in 1999, the only scientific studies cited were those published by psychologists in psychology journals. Today, psychology is engaged in an active dialogue with judges, police, and prosecutors on ways to improve the probative (evidentiary) value of eyewitness reports. The credibility of scientific psychology has risen immensely in the legal system recently, largely because psychologists were already "blowing the whistle" on eyewitness evidence well before forensic DNA testing began uncovering mistaken identifications in the 1990s. In effect, psychologists were able to use experiments to identify eyewitness problems long before the legal system was smacked in the face with DNA exonerations.

A primary purpose of this article is to describe empirical evidence supporting the proposition that some of the problems with eyewitness evidence can be addressed by improving the way the evidence is collected and preserved. We discuss how eyewitnesses are interviewed, how lineups are conducted, and why procedures can have a strong impact on the resulting probative value of eyewitness testimony. These variables are called system variables, because they are under the control of the justice system (Wells, 1978). The importance of system variables that can reduce eyewitness error has become increasingly apparent in light of the proven inadequacies of traditional safeguards against eyewitness mistakes, such as the presence of counsel at lineups and the opportunity to present motions to suppress suggestive procedures (Stinson, Devenport, Cutler, & Kravitz, 1996, 1997). But even if the system reaches a point at which it makes perfect use of system variables, eyewitness errors attributable to other factors will remain. Thus, it is important to review these other (non-system-controlled) factors as well.

This monograph is not intended as an exhaustive review of the eyewitness literature. Instead, we focus on practices, procedures, and research that address the most common threats to eyewitness reliability. Although the bulk of the scientific and legal literature we cite has a North American origin, the international research community has made extremely important contributions. In fact, historically, it was Europeans who played the much greater role in the study of eyewitness memory.[1]

We begin with a brief history of psychology's attempt to help the legal system on the eyewitness issue. Then we describe the DNA exoneration cases that began to unfold in the 1990s and the role these exonerations have played in giving scientific psychology a stronger voice in the legal system's policies and procedures involving eyewitness evidence. We then give an overview of the standard methods used in eyewitness research, followed by selected findings on estimator and system variables.

[1]We are fortunate to have Siegfried Sporer, a strong European contributor to the empirical literature on eyewitness issues, write the editorial preceding this monograph (see p. i). Sporer places our report in a broader historical and international context.

A BRIEF HISTORY OF EYEWITNESS PSYCHOLOGY

In his book *La Suggestibilité*, Alfred Binet (1900) argued for the creation of a practical science of testimony based on his observations about the effects of suggestion. Binet was the first to report that suggestive questioning influenced responses. But it was German psychologists who were among the first to argue that how eyewitnesses were questioned makes a great deal of difference. Louis William Stern was publishing and editing studies of eyewitness testimony as early as 1904 (Stern, 1904). In the United States, Guy Montrose Whipple published a number of articles in *Psychological Bulletin* on eyewitness testimony (Whipple, 1909, 1910, 1911, 1912). But it was Hugo Münsterberg's (1908) book *On the Witness Stand* and his injection of himself into the legal system that had a more lasting impact in the United States.

Münsterberg was recruited by William James in 1892 to come to Harvard to run the university's psychological laboratory. Münsterberg was very much a public figure and he appeared frequently in the popular press. He also was a somewhat controversial figure at Harvard, presumably because his colleagues did not see a great deal of merit in applying psychology. His lectures and writings were extremely perceptive and well reasoned, albeit rather short on data by modern standards. His prescience is evident in such matters as his claim that eyewitness certainty has a tenuous relation to accuracy and that while jurors might understand forgetting, they are not likely to understand that a witness can remember the wrong thing.

Although Münsterberg maintained a certain prominence in psychology, his impact on the legal system was muted dramatically by the skilled counterargumentation of one of the greatest minds in American jurisprudence, John Henry Wigmore. Particularly problematic for Münsterberg was a law review article by Wigmore (1909) that challenged Münsterberg's (1908) overstatements about the ability of psychology to help the legal system. Wigmore was especially effective in arguing that psychology did not yet have ready tools for handling the problem of evaluating eyewitness accounts, as Münsterberg had claimed. For the most part, Wigmore won the argument, at least from the perspective of the legal system.

Eyewitness research fell to a trickle in the period of the 1920s to 1960s. Some important work was done in the 1930s by Burtt (1931) and Stern (1939). The 1940s produced some important studies by Snee and Lush (1941) on question effects and by Allport and Postman (1947) on person-to-person information transfer. And although Hastorf and Cantrill (1954) demonstrated the effects of personal prejudice on perception in the 1950s, there was little discussion of the relevance of this to the legal system and to eyewitnesses in general. There are differing accounts of why these decades were largely devoid of eyewitness psychology. Sporer (1982) argues that it was the result of zealous overgeneralizations by psychologists that failed to meet the needs and standards of the courtroom.

The Modern Era of Eyewitness Research

More than any other individual's work, it was Elizabeth Loftus's elegant experiments on postevent information that gave rise to the modern era of eyewitness research. Loftus managed to show that realistic stimuli, such as pictures of stop signs and red barns in their natural settings, could be used in rigorous scientific experiments that revealed basic phenomena in memory and also had practical utility for understanding eyewitness error. By publishing her work in prestigious scientific psychology journals in the mid- and late 1970s—journals such as *Cognitive Psychology*, *Journal of Verbal Learning and Verbal Behavior*, and *Journal of Experimental Psychology: Human Learning and Memory*—Loftus legitimized the study of eyewitnesses in the minds of psychological scientists. Her book *Eyewitness Testimony* (Loftus, 1979) remains one of the best known psychology books almost three decades after it was released. Like Münsterberg, Loftus was criticized for some of her claims (e.g., McCloskey & Egeth, 1983), but, unlike Münsterberg, she helped spawn a new generation of researchers who have carefully and strategically built an empirical literature that the legal system must contend with.

While Loftus was focusing on memory for events and the malleability of memory, Robert Buckhout at Brooklyn College was focusing on memory for people. Buckhout was more concerned with mistaken identification from lineups than with memory for objects. Although Buckhout wrote a highly visible article in *Scientific American* reviewing research on eyewitness reliability (Buckhout, 1974), he was not otherwise particularly successful in getting his work published in scientific psychology journals. He did, however, create his own "in house" outlet called *Social Action and the Law*. Buckhout often used dramatic means to get his point across. For example, he got a New York City television station to broadcast a staged mugging followed by a six-person lineup. Of the 2,145 viewers who called in, nearly 2,000 mistakenly identified the mugger in the lineup (Buckhout, 1980). It is possible that Buckhout could have published some of his experiments in better journals but chose not to spend the time and effort required to go through the rigorous review process. Still, Buckhout influenced many younger researchers, who took up the issue of mistaken identification. At about the same time, eyewitness research activity was growing in the United Kingdom, prompted by the investigation of the Devlin Committee (Devlin, 1976; see also Bull & Clifford, 1976; Clifford & Bull, 1978; Davies, Ellis, & Shepherd, 1978; Ellis, Davies, & Shepherd, 1977).

One of the organizing themes that emerged from the 1970s was the distinction between system variables and estimator variables (Wells 1978). The argument was that some of the variables that affect the accuracy of eyewitness reports were under the control (or potentially under the control) of the justice system (system variables) while others were not (estimator variables). For example, how eyewitnesses are interviewed by police and how eyewitnesses are instructed prior to viewing a lineup are system variables, because they can be controlled by the system that is collecting the eyewitness evidence. Other variables—such as cross-race versus within-race identifications or stress experienced by the witness during the event—cannot be controlled by the system. Both system and estimator variables can be controlled and manipulated in experiments, but only system variables can be controlled in actual cases. Variables that cannot be controlled by the justice system (even though they can be controlled in experiments) are called estimator variables because the best that eyewitness psychology can do is help estimate their impact in a given case.

The singling out of system variables was important, because it addressed the primary argument that Wigmore used in his devastating criticism of Münsterberg—namely, that psychology had no practical recommendations for dealing with the eyewitness problem. Loftus's main findings fit nicely into the system-variable framework. For instance, if certain types of questions (leading questions) result in eyewitnesses incorporating information into their later reports regarding matters they did not witness, then psychology could devise practical ways to avoid this problem. Likewise, if certain instructions to eyewitnesses prior to viewing a lineup reduce the chances of mistaken identification, then psychology could advise on the best ways to instruct eyewitnesses.

Throughout the 1970s and 1980s, eyewitness research was largely ignored by the criminal justice system. The big exception was criminal defense lawyers. Defense lawyers were quick to recognize the potential for psychology to help them convince juries that eyewitness memory was not to be trusted, and they saw expert testimony as the mechanism to do this. The battle to permit expert testimony on eyewitness issues, however, was and is a contentious one. Expert testimony has been both permitted and denied in nearly every state in the United States, depending on the discretion of the trial judge. Prosecutors generally use four arguments against the admission of expert testimony on eyewitness issues. One argument is that the eyewitness literature is not sufficiently mature or precise to be considered scientific. Today, this argument almost never prevails. However, the three other arguments continue to prevent expert testimony on eyewitness issues in many jurisdictions. One is that such testimony invades the province of the jury, because it is the jury that must decide the credibility of witnesses. Another argument is that the findings are merely a matter of common sense and that juries already know these things from their everyday experience. Yet another argument is that the prejudicial value of expert testimony regarding eyewitnesses outweighs its probative value. This argument assumes that eyewitness experts can make juries more dubious of the eyewitness than they ought to be. It is not the purpose of the current monograph to argue the merits of expert testimony. We simply note that expert testimony for the defense was, until recently, virtually the only way the legal system acknowledged the scientific study of eyewitnesses.

Forensic DNA Testing: An Awakening of the Legal System

Much has changed in the past few years, but not because of any change in how eyewitness scientists have approached their work. Rather, the advent of forensic DNA testing has changed the way the legal system views eyewitness evidence. Previous studies of the conviction of innocent people had shown that mistaken eyewitness identification was implicated in the majority of wrongful conviction cases (e.g., Borchard, 1932; Frank & Frank, 1957; Huff, Rattner, & Sagarin, 1986). But it was the development of forensic DNA testing in the 1990s that permitted definitive cases of the conviction of innocent people in the United States to be uncovered. Defense lawyers Barry Scheck and Peter Neufeld, cofounders of the Innocence Project in New York City, took the lead and are still the central figures in facilitating the use of forensic DNA to test claims of innocence by people who were convicted by juries. Scheck and Neufeld were quick to see the pattern: Eyewitness-identification error was at the heart of the evidence used to convict the vast majority of these innocent people. Press accounts of these exonerations caught the attention of U.S. Attorney General Janet Reno, and an early report commissioned by Reno revealed that 26 of the first 28 exonerations were cases of mistaken eyewitness identification (Connors, Lundregan, Miller, & McEwan, 1996). Follow-ups revealed that 36 of the first 40 DNA exonerations were mistaken-identification cases (Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998). Scheck, Neufeld, and Dwyer (2000) reported that 52 of the first 62 DNA exonerations were mistaken-identification cases. As of this writing, there have been more than 180 definitive DNA exonerations; the proportion that involves mistaken eyewitness identification continues to run about 75% or more. The Innocence Project in New York maintains an up-to-date Web site, www.innocenceproject.org, that catalogues these DNA exonerations, and there are now innocence projects worldwide (http://forejustice.org/wc/wrongful_conviction_websites.htm).

Before the DNA exoneration cases, some people believed that the results of eyewitness experiments in psychology were mere academic exercises, games played with people's memories that would not apply to real witnesses and real crimes. At the very least, the DNA exonerations have proved that eyewitnesses can be absolutely positive and yet absolutely mistaken, just as was found in the experiments. But do 180-plus cases of mistaken identification prove anything? If these cases were the total, then it might be argued that this is a rather small fraction of convictions. But consider the following observations. Virtually all of these DNA exoneration cases involved sexual assault. Some also involved murder, robbery, and other offenses, but sexual assault is the common feature.

It is not that sexual assault witnesses are especially poor eyewitnesses. In fact, they might be the very best at identifying their attackers, because they tend to get longer, closer views of them than do victims of most other crimes. The reason these DNA exoneration cases are sexual assault cases is because they are the cases for which biologically rich DNA traces were left behind by the perpetrator in the form of semen. (In 2004, nearly 95,000 sexual assaults were reported, with a 43% clearance rate. For crime statistics, see www.fbi.gov/ucr/cius_04/offenses_reported/violent_crime/index.html.) Stranger-rape cases, in which identification is most likely to be an issue, constitute less than a third of all reported sexual assaults. More than 70% of reported sexual assaults involve an intimate partner, relative, or acquaintance, so about 30,000 cases of stranger-rape come to the attention of the police each year. In contrast to sexual assault cases, only a small fraction of murders (more than 16,000 reported in 2004) and almost no robberies (more than 400,000 reported in 2004) or aggravated assaults (more than 850,000 reported in 2004) result in biologically rich trace evidence being left behind. What can the person who was convicted of a convenience store robbery or a drive-by shooting use to prove that the eyewitness identification was mistaken? Thus, these 180-plus DNA exonerations represent a small proportion of the crimes for which eyewitness identification evidence has been used to convict people. Furthermore, only a fraction of old sexual assault convictions can now be tested, because the evidence was never collected, was collected improperly, has deteriorated, has been lost, or has been destroyed. All in all, the 180 (and growing) DNA exonerations can only be a small fraction of the total number of cases in which people have been convicted because they were mistakenly identified by eyewitnesses.

We will not venture an estimate of the number of people in prison who are innocent victims of mistaken eyewitness identification. Instead, our focus is on what the legal system might be able to do to help prevent these mistakes from occurring in the future. This is where we have seen some promising progress recently. Janet Reno's appointment of a working group to develop guidelines for eyewitness evidence was a watershed event, because the group included five eyewitness researchers. Reno recognized that scientific psychology was well ahead of the legal system both in recognizing the eyewitness problem and in developing solutions for it. An account of this process, which yielded the first set of U.S. national guidelines on eyewitness evidence, has been published elsewhere (Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000). Since the publication of the guide, a number of jurisdictions have formally adopted the recommendations and have gone well beyond the guide to include procedural changes recommended by eyewitness scientists. These jurisdictions include the states of New Jersey, North Carolina, and Wisconsin, as well as the cities of Boston and Minneapolis, among others (Wells, 2006).

Despite these encouraging reforms, it is estimated that only about 10% of the U.S. population reside in reformed jurisdictions (Wells, 2006). Will these system-variable improvements continue by increasing numbers of jurisdictions in the years to come? Only time will tell. In the following sections we review some of the evidence that has led to the changes, and we note how the eyewitness-research area must continue to develop to ensure that the evolving relationship between the legal system and psychological science will be a fruitful and lasting one.

COMMON METHODS USED IN EYEWITNESS RESEARCH

The experimental method has dominated the eyewitness literature, and most of the experiments are lab based. Lab-based experimental methods for studying eyewitness issues have strengths and weaknesses. The primary strength of experimental methods is that they are proficient at establishing cause–effect relations. This is especially important for research on system variables, because one needs to know in fact whether a particular system manipulation is expected to cause better or worse performance. In the real world, many variables can operate at the same time and in interaction with one another. Multicollinearity can be quite a problem in archival/field research, because it can be very difficult to sort out which (correlated) variables are really responsible for observed effects. The control of variables that is possible in experimental research can bring clarity to causal relationships that are obscured in archival research. For example, experiments on stress during witnessing have shown, quite compellingly, that stress interferes with the ability of eyewitnesses to identify a central person in a stressful situation (Morgan et al., 2004; Deffenbacher, Bornstein, Penrod, & McGorty, 2004). However, when Yuille and Cutshall (1986) studied multiple witnesses to an actual shooting, they found that those who reported higher stress had better memories for details than did those who reported lower stress. Why the different results? In the experimental setting, stress was manipulated while other factors were held constant; in the actual shooting, those who were closer to the incident reported higher levels of stress (presumably because of their proximity) but also had a better view. Thus, in the actual case, stress and view covaried.

The experimental method is not well suited to postdiction with estimator variables—that is, there may be limits to generalizing from experiments to actual cases. One reason is that levels of estimator variables in experiments are fixed and not necessarily fully representative of the values observed in actual cases. In addition, it is not possible to include all interesting and plausible interactions among variables in any single experiment (or even in a modest number of experiments). Clearly, generalizations to actual cases are best undertaken on the basis of a substantial body of experimental research conducted across a wide variety of conditions and employing a wide variety of variables. Nevertheless, the literature is largely based on experiments due to a clear preference by eyewitness researchers to learn about cause and effect. Furthermore, "ground truth" (the actual facts of the witnessed event) is readily established in experiments, because the witnessed events are creations of the experimenters. That kind of ground truth is difficult, if not impossible, to establish when analyzing actual cases.

Experimental Methods

The ecological validity of witnessed events (when examined at the surface level) varies considerably across experiments. Some eyewitness experiments simply show slides to participant witnesses, while others stage live crimes. Some of the staged crimes have been elaborate ruses in which calls are made to "police" (actually confederates of the experimenter) and participants are shown lineups while still believing that what they witnessed was real (e.g., Luus & Wells, 1994; Wells & Murray, 1983). Perhaps the most common witnessed event used by researchers is the video crime. The immense time and cost involved in staging live crimes has undoubtedly made them less common in the literature in recent years, but the success of video crime experiments in the peer-review process suggests that researchers believe this method manages to capture the elements that are important for studying eyewitness processes. Usually, the memory-acquisition process is incidental in the sense that the participant witnesses do not know when they watch the video that the study concerns eyewitness memory. Instead, researchers commonly tell them that they are going to have to form impressions or make judgments about the people or scenes. Only later are they informed that the study concerns eyewitness memory.

In lineup experiments, the participant witnesses are usually tested with photo lineups rather than with live lineups. Again, the savings in cost and time are factors, but the use of photo lineups in experiments parallels their use in actual cases. In the United Kingdom, there has been a move toward the use of video lineups (Pike, Kemp, Towell, & Phillips, 1997; Valentine & Heaton, 1999). Although some jurisdictions (such as New York) still use live lineups, most jurisdictions in the United States use photo lineups. Even where live lineups are in common use, more often than not they are preceded by a photo lineup, and the live lineup is merely a confirmatory tool. Thus, the prevalence of photo lineups in experiments reflects what is happening in actual criminal investigations.

It is standard practice in experiments to use lineups in which the actual perpetrator is present in the lineup for some participant witnesses and not present for others. The not-present lineups (target-absent or perpetrator-absent lineups) are critically important for eyewitness-identification studies that are designed to examine accuracy. Target-absent lineups simulate the real-world situation in which police have focused their suspicion on an innocent suspect. The standard procedure in lineup experiments is to create a target-absent lineup by replacing the target with another person who fits the target's description and leaving the fillers (the innocent distracters or foils in the lineup) the same.

Participant witnesses in experiments typically take the perspective of a bystander rather than a victim. However, some experiments have examined possible differences between bystander eyewitnesses and victim eyewitnesses and have found no significant differences (Hosch & Cooper, 1982; Hosch, Leippe, Marchioni, & Cooper, 1984).

Participant witnesses in experiments are typically college students. The reliance on this population has been criticized, especially by prosecutors. However, many experiments have included other populations, such as young children, adults, and the elderly. Importantly, when differences are found, the results favor the college students. Specifically, college students are less suggestible and more accurate as eyewitnesses overall than are either children or the elderly (Cutler & Penrod, 1995; Searcy, Bartlett, & Memon, 1999). Presumably this is due to the higher education level, intelligence, memory ability, visual acuity, alertness, and general health of college students relative to the general population. Thus, if anything, heavy reliance on college student subject populations for eyewitness research may paint an unrealistically rosy picture of eyewitness abilities.

Within the basic eyewitness-experiment paradigm, manipulations are embedded and their effects are observed. For example, an experiment focusing on system variables might have everyone view the same simulated crime and then randomly assign some participant witnesses to receive a postevent suggestion or randomly assign some to receive a particular prelineup instruction. In an experiment focusing on estimator variables, participants might be randomly assigned to view a crime in which the perpetrator is of a different race or the same race or to make an identification after a short delay or after a long delay.

Archival Methods

Although the experimental method is preferred, archival studies of eyewitnesses have become more common in recent years. A major drawback to archival studies is the inability to establish cause and effect and the questionable basis for assuming ground truth. Studies of the DNA exoneration cases involve ground truth for identity of the perpetrator, but these are only case studies, not archival analyses. Archival analyses have proven to be particularly informative with regard to lineups. A lineup that is properly constructed includes only one suspect (who might or might not be the perpetrator); the other people in the lineup are innocent fillers who would not be charged with the crime if they were identified by the eyewitness. Thus, when an eyewitness selects a filler in an actual lineup, it is immediately classifiable as an error. It is not the type of error that could send an innocent person to jail (only identifications of an innocent *suspect* could do that), but it is an identification error nevertheless.

Archival analyses of filler identifications have yielded very interesting results. Wright and McDaid (1996) analyzed 1,561 lineup outcomes in London and found filler-identification rates of 19.9%. These data are similar to the 21% filler identification rate reported by Slater (1994) in a study of 843 lineups conducted by the Metropolitan Police in London. Behrman and Davey (2001) reported that 24% of identifications from live lineups in Sacramento, California, were identifications of fillers. Valentine, Pickering, and Darling (2003) analyzed 119 lineups in the greater London area and found that 21.6% of the eyewitnesses identified fillers. In these four studies of actual eyewitnesses to serious crimes, filler identifications constituted approximately one third of all positive identifications. These archival results represent a very important complement to the experimental studies of eyewitnesses for several reasons. First, they indicate filler-identification results that are quite consistent with rates obtained in experiments (Ebbeson & Flowe, n.d.; Steblay, Dysart, Fulero, & Lindsay, 2001). Second, these archival results address a common criticism of experiments—namely, that participant witnesses in experiments are not as cautious as actual crime witnesses are, because the consequences of a mistaken identification in an experiment are not serious. But the witnesses in the archival studies were actual witnesses to crimes and yet mistakenly identified fillers in one third of their positive identifications. Third, the filler-identification rates in the archival studies permit us to make conservative estimates of the risk that an innocent suspect would face in these lineups. For example, with five fillers in each lineup (six-person lineup minus the suspect) and a 20% filler-identification rate, the risk to any given filler is 4%. If an innocent suspect has the same risk as a filler, the estimated risk to an innocent suspect is 4%.

These estimates of the risk to an innocent suspect are conservative for two reasons. First, lineups rarely yield equal distributions of error because the innocent suspect will commonly stand out for any number of reasons, including the selection of fillers that bear a poor resemblance to the description of the perpetrator given by the witness (Valentine & Heaton, 1999; Brigham, Meissner, & Wasserman, 1999). Second, when the actual perpetrator is not in the lineup (i.e., the suspect is innocent), the rates of filler identification increase (see Wells & Olson, 2002). Assuming that the perpetrator was present in a large proportion of the lineups in these archival studies, the filler-identification rates underestimate the expected error rate for any given lineup in which the perpetrator is absent.

Archival studies also permit analyses that examine results as a function of different levels of critical variables. For example, Wright and McDaid (1996) found that the filler-identification rate was 20.8% for violent crimes and 17.6% for nonviolent crimes. Valentine et al. (2003) found that the filler-identification rate was 15.9% when a weapon was present and 23.7% when there was no weapon. The latter result seems peculiar in light of the experimental results indicating a deleterious effect for the presence of a weapon (see meta-analysis by Steblay, 1992)—but in the weapons-effect section later in this monograph, we note that archival data are subject to "selection effects" that may offset or reinforce the effects of variables such as weapon focus.

Another interesting archival finding does not concern eyewitnesses per se but has a powerful bearing on expected rates of mistaken identification in the courtroom: Archival studies indicate that those charged with a crime enter a guilty plea in 80 to 90% of cases (Cole, 1986). Let us assume that 80% plead guilty (the argument is stronger at 90%). We might assume that no mistakenly identified (innocent) suspects plead guilty and that all the guilty pleas are from guilty suspects. (In no sense do we intend for this assumption to be interpreted as a denial of the important work of Kassin & Gudjonsson, 2004, and other false-confession researchers, who have clearly made a compelling case that innocent people plead guilty.) Even if we presume that 10% of mistakenly identified suspects plead guilty, 90% of the innocent suspects and only 20% of the guilty suspects will go to trial. Assume further that a mere 4% of suspects identified from a lineup are innocent and 96% are guilty. If we assume that 80% of guilty suspects plead guilty and therefore do not go to trial, only 20% of the 96% (19.2% of the guilty) will go to trial, whereas 90% of the 4% (3.6% of the innocent suspects) will go to trial. Thus, at the trial level, 16% of the defendants (3.6% of the 22.8% going to trial) will be cases of mistaken identification. Charman and Wells (2006) called this the "pleading effect"; it illustrates how the mistaken-identification rate can be expected to be higher at the trial level than at the lineup level (see Fig. 1).

ESTIMATOR VARIABLES

We first review estimator variables. Although these variables are not under the control of the justice system, they are important to our treatment for two main reasons. First, estimator variables are central to our understanding of when and why eyewitnesses are most likely to make errors. Informing police, prosecutors, judges, and juries about the conditions that can affect the accuracy of an eyewitness account is important. Second, our understanding of the importance of any given system variable is, at least at the extreme, dependent on levels of the estimator variables. Consider a case in which a victim eyewitness is abducted and held for 48 hours by an unmasked perpetrator; the witness has repeated viewings of the perpetrator, lighting is good, and so on. We have every reason to believe that this witness has a deep and lasting memory of the perpetrator's face. Then, within hours of being released, the eyewitness views a lineup. Under these conditions, we would not expect system variables to have much impact. For instance, a lineup that is biased against an innocent suspect is not likely to lead this eyewitness to choose the innocent person, because her memory is too strong to be influenced by lineup bias. On the other hand, when an eyewitness's memory is weaker, system variables have a stronger impact.

The effects on identification accuracy of a large number of estimator variables—witness, crime, and perpetrator characteristics—have been investigated by psychologists. Here we recount findings concerning several variables that have received significant research attention and achieved high levels of consensus among experts (based on items represented in a survey by Kassin, Tubb, Hosch, & Memon, 2001) or have been the subject of interesting recent research.

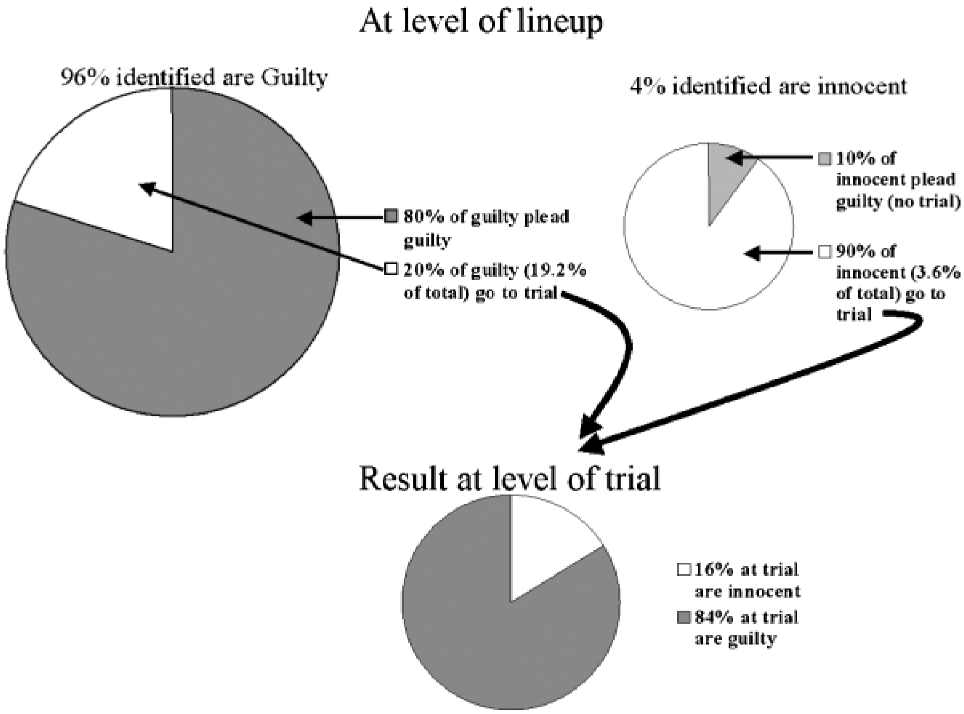

Fig. 1. The ''pleading effect'' (Charman & Wells, 2006). Assuming first that 96% of suspects identified from a lineup are guilty and 4% are innocent, if 80% of the guilty suspects and 10% of the innocent suspects plead guilty (thereby foregoing a trial), the result is that 16% of the defendants who go to trial will be innocent—that is, cases of mistaken identification.

Cross-Race Identification

Meissner and Brigham (2001a) published the most recent broad review of research on the problems associated with what has sometimes been called other-race or cross-race identification impairment or own-race bias (ORB). Meissner and Brigham analyzed data from 39 research articles, with 91 independent samples involving nearly 5,000 participant witnesses. They examined measures of correct identification and false-alarm rates, as well as aggregate measures of discrimination accuracy and response criterion. They reported that the chance of a mistaken identification is 1.56 times greater in other-race than in same-race conditions and that the witnesses were 1.4 times more likely to correctly identify a previously viewed own-race face as they were to identify an other-race face. Participants were more than 2.2 times as likely to accurately categorize own-race faces as new versus previously viewed as they were to accurately categorize other-race faces. Meissner and Brigham explored the question of whether cross-race contact might reduce these effects and found that such contact played only a small role in ORB, accounting for just 2% of the variability across participants (see also Wright, Boyd, & Tredoux, 2003). They also found that the amount of viewing time available to witnesses significantly influenced ORB; specifically, false alarms to other-race faces increased when study time was limited.

Recent research by Pezdek, Blandon-Gitlin, and Moore (2003) examined cross-race impairment in kindergarten children, third graders, and young adults who viewed black and white target faces and a day later were tested with a six-person lineup. These researchers observed the usual cross-race effect, which did not differ across age groups: In each age group, cross-race identification was less accurate than own-race identification.

Stress

Despite the importance of knowledge about the effects of stress on witnesses, researchers cannot simulate violent crimes and pose a threat to the well-being of experimental subjects. Researchers have therefore resorted to a variety of manipulations to induce stress, including the use of violent versus nonviolent videotaped crimes. Increased violence in videotaped reenactments of crimes has been shown to lead to decrements in both identification accuracy and eyewitness recall (Clifford & Hollin, 1981; Clifford & Scott, 1978), but this finding is not universal (Cutler, Penrod, & Martens, 1987a).

Deffenbacher et. al (2004) recently published a meta-analysis of stress effect studies. The meta-analysis was conducted on 27 tests of the effects of heightened stress on identification accuracy and on 36 tests of its effect on recall of crime-related details. They found that high levels of stress negatively affected both types of memory. The effect of stress was notably larger for target-present than for target-absent lineups—that is, stress particularly reduced correct identification rates. The effect was also

TABLE 1

Percentages of Accurate and Mistaken Identifications From Study of Eyewitness Identification Under High Versus Low Stress

	High stress	Low stress
Correct identifications [target-present]		
Live lineup method	27	62
Photospread method	36	76
Sequential photo method	49	75
Mistaken identifications [target-absent]		
Live lineup method	45	50
Photospread method	48	61
Sequential photo method	0	0

Note. Source: Morgan et al. (2004).

considerably larger for eyewitness-identification studies that simulated eyewitness conditions (e.g., staged crimes) than for face-recognition studies.

These effects are well illustrated in a study by Morgan et al. (2004) that examined the eyewitness capabilities of more than 500 active-duty military personnel enrolled in a survival-school program (see Table 1). After 12 hours of confinement in a mock prisoner-of-war camp, participants experienced both a high-stress interrogation with real physical confrontation and a low-stress interrogation without physical confrontation. Both interrogations were 40 minutes long; they were conducted by different persons. A day after release from the camp, and having recovered from food and sleep deprivation, the participants viewed a 15-person live lineup, a 16-person photo spread, or a sequential presentation of photos of up to 16 persons. Regardless of the testing method, as Table 1 shows, memory accuracy for the high-stress interrogator was much lower overall than for the low-stress interrogator.

Weapon Focus

Weapon focus refers to the visual attention eyewitnesses give to a perpetrator's weapon during the course of a crime. It is expected that the attention the eyewitness focuses on the weapon will reduce his or her ability to later recall details about the perpetrator or to recognize the perpetrator. Researchers have assessed eyewitness recall of various crime details in an attempt to establish the parameters of weapon-focus effects on perception and memory; these efforts were reviewed in a meta-analysis by Steblay (1992). The review included 19 studies with a total sample of 2,082 participants. The weapon-focus effect on identifications was statistically significant but reflected a modest impairment; the effect on description accuracy was larger. The analysis indicated that the weapon-focus effect was larger in target-absent lineups and when memory was generally impaired. Research by Mitchell, Livosky, and Mather (1998); Pickel (1998, 1999); and Shaw and Skolnick (1999) indicates that any surprising object can draw attention away from the perpetrator and that novelty, rather than threat, may be the critical ingredient in the effect.

Researchers have tried to detect weapon-focus effects in field studies, and the results are somewhat conflicting. Tollestrup, Turtle, and Yuille (1994) examined the effect of weapon focus on the rate of suspect identification and obtained data consistent with laboratory findings. But Valentine et al. (2003) did not find a weapon-focus effect in their study of 640 attempts by eyewitnesses to identify the alleged target in 314 lineups. Of course, as noted earlier, in nonexperimental studies it is difficult to control for variables that might obscure a weapon-focus effect. For example, in the study by Valentine et al., the primary outcome variable is suspect choices rather than perpetrator choices (i.e., witness identifications are intended to determine whether suspects are perpetrators)—whereas in experimental research the identity of the perpetrator is known to the researcher.

Field research can also suffer from *selection effects* that can obscure the effects of variables of interest. For example, a true weapon-focus effect could be obscured if witnesses to crimes involving weapons believe that their memory is weak and are therefore less inclined to attend lineups. The result could be a reduction in the number of weapon-focus-impaired witnesses presented with lineups and thus a reduced number of cases of weapon focus.

As mentioned earlier, a selection effect might actually reduce our concern about the potential impact of weapon focus on eyewitness performance. On the other hand, it is conceivable that more intensive police investigations of weapon-present cases produce a higher proportion of perpetrator-present lineups for weapon-present witnesses, with the result that the apparent performance of weapon-present witnesses is improved even though their memories are impaired. If investigations of all crimes were similarly intense, a weapon-focus effect might emerge. One might also imagine that the police are more motivated to "help" weapon-present witnesses identify perpetrators who use weapons and who thus pose a threat to society. Such help might take the form of suggestive instructions to witnesses and suggestive lineups.

Exposure Duration

Common sense tells us that the amount of time available for viewing a perpetrator is positively associated with the witness's ability to subsequently identify him or her. A meta-analysis by Shapiro and Penrod (1986) showed that the linear trend for exposure time was associated with improved performance. The effects of exposure time were illustrated in a study by Memon, Hope, and Bull (2003) in which mock witnesses viewed a realistic videotaped crime in which the target/perpetrator was visible for 12 versus 45 seconds. Witnesses were tested with target-present and target-absent arrays 40 minutes later. The proportion of correct identifications in target-present arrays and correct rejections in target-absent arrays increased substantially when exposure time increased from 12 seconds to 45 seconds (from 32% to 90% for correct identifications and from 15% to 59% for correct rejections), although mistaken identifications in target-absent arrays remained high even with longer exposure (85% at 12 seconds and 41% at 45 seconds).

Disguise

It is common for people to don disguises before engaging in criminal acts. Full-face masks, stockings, hats, and hoods can be quite effective in diminishing the facial-feature cues necessary for recognition (Cutler, Penrod, & Martens, 1987a, 1987b; McKelvie, 1988; Patterson & Baddeley, 1977). For example, Cutler et al. (1987b) had participants view a videotaped liquor store robbery and later attempt an identification from a videotaped lineup. In half of the robberies, the robber wore a knit pullover cap that covered his hair and hairline. In the other half, he did not wear a hat. The robber was less accurately identified when he was disguised: 45% of the participants identified the robber in the lineup test if he wore no hat during the robbery; only 27% identified him if he wore a hat during the robbery.

Shapiro and Penrod, in their 1986 meta-analysis, coded experiments for whether or not faces were changed between the initial viewing and recognition phases. Transformations included changes in facial hair and deliberate disguises, such as masks or hats. Nontransformed faces were more accurately recognized (effect size $d = 1.05$; 75% vs. 54%) and less often falsely identified ($d = .40$; 22% vs. 30%) than transformed faces were.

Not all disguises or changes in appearance work. Yarmey (2004) found similar levels of identification accuracy for a young woman viewed for 15 seconds in naturalistic circumstances, regardless of whether or not she wore a baseball cap and dark sunglasses. There was, however, an interaction involving disguise: Witnesses who were given enhanced retrieval instructions (involving mental rehearsal of the encounter) made significantly more correct rejections in the no-disguise condition than in the disguise condition.

Retention Interval

Common sense tells us that memory declines over time. Can we expect eyewitness-identification accuracy to decline as the time between the crime and the identification test increases? Shapiro and Penrod (1986) included retention interval in their meta-analysis. When studies that manipulated retention interval were grouped into long versus short time delays (the exact manipulation depended on the study), longer delays led to fewer correct identifications ($d = .43$; 51% vs. 61%) and more false identifications ($d = .33$; 32% vs. 24%). Across all the studies examined in that meta-analysis (including those that did not directly manipulate retention interval), retention interval also proved an important determinant of correct identifications ($r = -.11, p < .05$), although there was no significant relationship with false identifications.

Witness Intoxication

Read, Yuille, and Tollestrup (1992, Experiment 1) tested identification accuracy one week after a staged event using a six-person lineup; they found that alcohol intoxication while witnessing the event was associated with a lower rate of correct identifications when the level of arousal (manipulated by varying the participants' perceptions of the probability of getting caught stealing an item from an office) was low during the event. False identification rates were the same for intoxicated and sober participants. Of course, after one week the participants were no longer intoxicated, which raises the question of what the effect of intoxication at viewing and identification would be.

Dysart, Lindsay, MacDonald, and Wicke (2002) note that the popular belief is that intoxicated witnesses are less accurate than sober witnesses. However, one theory concerning "alcohol myopia" (Steele & Josephs, 1990) predicts an interaction between blood-alcohol level and identification procedures in which witnesses who were intoxicated at encoding will be less accurate only in target-absent conditions. The theory suggests that, compared with intoxicated witnesses, sober witnesses will encode more information/cues about the perpetrator, which will facilitate correct rejections in target-absent procedures. Intoxicated witnesses are likely to encode only salient cues, and erroneous identifications will result where more subtle cues would have indicated that the suspect was not the target. On the other hand, using salient cues will be effective for intoxicated witnesses when the target is present.

Dysart et al. (2002) examined the effect of alcohol consumption on identification accuracy using "showups," a procedure in which the witness is shown the suspect alone, without any fillers. A showup is the identification procedure most likely to be used by police with intoxicated witnesses. As predicted, the researchers found that in the target-present showup condition, blood-alcohol level was not significantly related to correct identification; however, in the target-absent condition, higher blood-alcohol levels were associated with a higher likelihood (52%) of a false identification than were lower blood-alcohol levels (22%).

SYSTEM VARIABLES

System variables (variables that can be controlled in actual cases) tend to center on factors that come into play after the witnessed event has passed. At that point, the legal system has some control over a number of important variables, but not necessarily all variables. For instance, first responders at a crime scene can separate eyewitnesses so they do not influence each other, but some interactions could have already occurred before the arrival of investigators. Similarly, although investigators have total control over how a lineup is conducted, some identifications occur outside the control of the legal system—for example, when an eyewitness spontaneously identifies someone on the street as the perpetrator of an earlier crime.

System variables tend to be divided into two broad categories. One category is interviewing eyewitnesses, a process that generally involves recall memory. The other category is the identification of suspects, a process that generally involves recognition memory. It is important to note that neither interviewing nor identification is considered by eyewitness scientists to be purely a memory process. Social influence can be a huge factor in both.

The case of James Newsome, a man who served 15 years for a murder he did not commit, is an extreme example of an eyewitness making a positive identification from a lineup, even though his memory told him that the man he identified was not the man who committed the murder. After Newsome was proved innocent and the actual perpetrator was found through physical evidence, eyewitness Anthony Rounds came forward and described how Chicago police had forced him to identify Newsome from the lineup, even though he knew that Newsome was not the man he saw commit the murder. According to Rounds, the lineup administrators told Rounds whom to identify; when he resisted, their intimidating insistence led him to identify Newsome and give confident identification testimony at trial. A lawsuit in 2002 yielded strong evidence to support Rounds's claim, and a jury awarded damages to Newsome; the finding was upheld by the U.S. Seventh Circuit Court of Appeals (Newsome v. McCabe et al., 2002).

Although this is an extreme example, it illustrates how extraneous external variables can influence eyewitness testimony without operating through memory mechanisms. Under other circumstances, social-influence variables are thought to actually influence memory. For instance, a misleading question such as "What kind of hat was the gunman wearing?" when the gunman had no hat could lead an eyewitness to develop a memory for a hat that did not exist. For these reasons, eyewitness scientists concern themselves with both social-influence variables and memory variables.

Interviewing Eyewitnesses

Research on interviewing eyewitnesses dates back to the early 1900s. Alfred Binet (1900) was the first to study suggestibility in children in France, and William Stern (1904) initiated eyewitness research on interrogation in Germany. Snee and Lush (1941) wrote a short empirical article on the use of interrogatory versus narrative methods of interviewing eyewitnesses. Modern research on the issue undoubtedly owes much to the influence of Elizabeth Loftus, who used the method of asking questions of eyewitnesses to implant misleading information (e.g., Loftus & Palmer, 1974). This line of research paved the way for experimental studies of the effects of explicit and subtle forms of misinformation imparted during questioning of adult and child witnesses (for reviews see Bruck & Ceci, 1999; Loftus, 2005; Wright & Loftus, 1998). This work led to important theoretical advances in our understanding of the mechanisms underlying eyewitness suggestibility in interviews. Examples include the source-monitoring framework (Lindsay & Johnson, 1989; Mitchell & Johnson, 2000; Poole & Lindsay, 2001); fuzzy-trace theory (Brainerd & Reyna, 1998; Memon, Hope, Bartlett, & Bull, 2002); an activation-based memory model (Ayers & Reder, 1998); retrieval-induced forgetting (MacLeod, 2002); the role of metacognition (Koriat, Goldsmith, & Pansky, 2000); and the social-influence approach (Echterhoff, Hirst, & Hussy, 2005; Gabbert, Memon, & Wright, in press; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001).

In this monograph, we do not discuss the mechanisms responsible for distortions in information retrieved in eyewitness interviews. Instead, we use one example of a procedure that arose as a result of a direct request from the police to improve the probative value of eyewitness evidence. This example shows how researchers have attempted to wrap their knowledge about memory and social influence into a set of procedures for interviewing eyewitnesses. It is also the most developed and extensively researched procedural package for gathering detailed reports from cooperative eyewitnesses. (Readers who are interested in other approaches to interviewing eyewitnesses, including interviews designed to detect deception, should refer to reviews by Granhag & Stromwell, 2004; Memon & Bull, 1999; Poole & Lamb, 1998; and Vrij, 2000.)

The Cognitive Interview

The cognitive interview (CI) was initially developed by the psychologists R. Edward Geiselman (University of California, Los Angeles) and Ronald P. Fisher (Florida International University) in the early 1980s (Geiselman et al., 1984; Geiselman, Fisher, MacKinnon, & Holland, 1985) and has resulted in more than two decades of research. Two main forces drove the development of the CI. The first was a request from police officers and legal professionals to improve the practices of police interviewers when gathering information from eyewitnesses. Analysis of the techniques used by untrained police officers in Florida (Fisher, Geiselman, & Raymond, 1987) suggested that there existed some fundamental problems in the conduct of interviews, leading to ineffective communication and poor memory performance. The "standard police interview" was characterized by constant interruptions, excessive use of a predetermined list of questions with an expectation that witnesses could provide answers, and questions that were timed inappropriately. For example, if the witness was describing one of the perpetrators, the officer might switch the line of questioning to the actions of another perpetrator. Interestingly, the same problems were identified in studies of the typical police interview in Britain (George, 1991) and Germany (Berresheim & Weber, 2003).

TABLE 2
Revised Cognitive Interview Procedure

Step 1. Build rapport
- (a) Personalize the interview
  Exchange names. Make sure the witness is comfortable and is willing to try to remember as much as possible. Ask the witness to give as many details as possible but not to guess or fabricate.
- (b) Transfer control to the witness
  Tell the witness that you do not have knowledge of the event and it is the witness who holds all the relevant information. Let the witness choose the starting point for the narrative and give the account at his or her own speed and in his or her own words. Do not interrupt the witness, if at all possible. Listen actively to what he or she has to say. Allow for pauses.

Step 2. Recreate the context of the original event and ask the witness to report in detail.
  To reinstate context, invite the witness to close his or her eyes and place himself or herself back at the scene.

Step 3. Open-ended narration
- (i) Request narrative description
  Ask the witness to give a narrative account of the event in his or her own words. If clarification is required, use open-ended questions. Do not interrupt the narration to ask questions, although prompts such as "tell more" may be used. Avoid judgmental comments and closed (yes/no) questions.
- (ii) Focused retrieval
  This is not a technique but a general guideline to follow to help the witness concentrate on what he or she is describing by
  - using open-ended questions
  - allowing for long pauses
  - not interrupting the witness when he or she is speaking
- (iii) Extensive retrieval
  Encourage the witness to search through his or her memory more extensively by asking him or her to report details from a number of different perspectives and in different chronological orders.
- (iv) Witness-compatible questioning
  Time the questions appropriately so they are compatible with the witness's retrieval pattern rather than adhering to a protocol.

Step 4. Closure
  Be sure to leave time to brief the witness and let him or her know what might happen next. Exchange contact information and encourage the witness to get in touch if he or she remembers additional details.

Note. Adapted from Fisher and Geiselman (1992).

The CI in its present form represents the alliance of two fields of study: communication and cognition. The social-psychological concerns of managing a face-to-face interaction and communicating effectively with a witness were integrated with what psychologists knew about the way people remember things. The social aspects are embodied in what is referred to as a *structured interview*, which consists of a phased procedure (free report followed by open-ended questions) and incorporates techniques to facilitate communication. These techniques include rapport building, which is designed to increase the transfer of control from the interviewer to the witness, and the use of a questioning strategy guided by the witness's own free report rather than one that is based on a predefined protocol. The structured interview resembles the recommended procedure for conducting investigative interviews with witnesses and victims in many countries (see Poole & Lamb, 1998; Westcott, Davies, & Bull, 2002).

The original version of the CI was presented as a set of four specific cognitive techniques for improving eyewitness recall. Following a series of laboratory simulations and field research, the procedure was revised in 1992 (Fisher & Geiselman, 1992). The version of the CI that has subsequently evolved focuses heavily on communication techniques and social dynamics and is a procedure in which the cognitive and communication components work in tandem. Here we will focus primarily on research and practice relating to the revised CI (also referred to in the literature as the "enhanced" CI). For a summary of the revised CI procedure, see Table 2.

The revised CI comprises several phases during which the interviewer engages with and establishes rapport with the witness, asks the witness to provide a narrative account of the witnessed event, and then probes with questions relating to the details the witness has provided. Throughout the process, the interviewer interrupts as little as possible, allows the witness to dictate the subject matter and sequence of questioning, and listens actively to what the witness has to say. One of the primary aims of the CI is to facilitate the exchange of information between the witness and interviewer through effective communication.

The first task of the interviewer is to build rapport with the witness. This rapport serves two functions. First it puts the witness at ease, minimizing the discomfort and distress sometimes associated with sharing an intimate or fearful experience with a stranger. Second, there is some evidence that building rapport with open-ended questions can increase the accuracy of a child witness's report (Roberts, Lamb, & Sternberg, 2004). An important component of rapport building in the revised CI is for the interviewer to explicitly "transfer control" to the witness (see Table 2 for details).

The "cognitive" part of the CI relies on two theoretical principles. First, a retrieval cue is effective to the extent that there is an overlap between the encoded information and the retrieval cue. Reinstatement of the original encoding context increases the accessibility of stored information (Tulving & Thomson's encoding specificity hypothesis, 1973). Second, multiple trace theory (Bower, 1967)—which proposes that memories are made up of networks of associations rather than discrete and unconnected incidents—states that a memory can be cued by several means and that information not accessible with one technique may be accessible with another.

Having established rapport with the witness, the interviewer instructs the witness to mentally reconstruct the physical and personal contexts that existed at the time of the crime. Interviewers can help witnesses by asking them to form an image or impression of the environmental aspects of the original scene (e.g., the location of objects in a room); to comment on any emotional reactions and feelings (e.g., surprise, anger) at the time; and to describe any sounds, smells, and physical conditions (e.g., hot, humid, smoky) that were present. Occasionally a witness can be taken back to the scene of the crime. Once the witness has mentally reconstructed the context, the interviewer asks him or her to provide a detailed account of the event (the free narrative). To extend retrieval, the witness is asked to report all details, including partial or incomplete memories. To minimize editing, Fisher and Geiselman (1992) advised interviewers to instruct witnesses to report everything that comes to mind, even if it is trivial or out of chronological order. In addition to facilitating the recall of additional information, this technique may yield information that can be valuable in piecing together details from different witnesses to the same crime. Roberts and Higham (2002) obtained ratings of the forensic relevance of details elicited with the CI by asking police officers and prosecutors to rate the relevance of each detail to a criminal investigation/court proceeding. At best, only 50% of the information the CI elicited was deemed relevant by forensic experts. Most of the correct, forensically relevant details appeared in the free-narrative account (cf. Memon, Wark, Bull, & Köhnken, 1997).

Once the witness has provided an open-ended account, the CI interviewer can probe for details using open-ended questions and, when appropriate, can ask follow-up questions to clarify what the witness has said. It is imperative that interviewers remind witnesses that if they are unsure of an answer to a question, they should say so and not guess. Appropriate sequencing of the interviewer's questions (referred to as interviewee-compatible questioning) is critical. Each eyewitness will have a unique mental representation of the event, depending on the details or aspects of the event he or she attended to and the order in which events unfolded for him or her (Fisher & Schreiber, in press). The interviewer should be guided by the interviewee's pattern of recall rather than adhering to a rigid protocol or predetermined checklist. For example, if an interviewee is describing a suspect's face, this indicates that the mental image of the perpetrator's face is currently active and details about the face are accessible (Pecher, Zeelenberg, & Barsalou, 2003). At this point, the interviewer should ask questions relating to the suspect's appearance and not switch to another topic, such as the suspect's car.

In a CI, the witness is encouraged to focus or concentrate on mental images of the various parts of the event, such as the suspect's face (Fisher & Geiselman, 1992). The interviewer exhausts the content of each image by asking the witness to form an image and then describe it in as much detail as possible. Bekerian and Dennett (1997) refer to this focus on specific features as "molecular imaging," as compared to the general "molar" approach, which emphasizes reinstating environmental context. To effectively engage the interviewee in focused retrieval, the interviewer must speak slowly and clearly, pausing at appropriate points to allow the interviewee time to create an image and respond (Memon, 2006). Unfortunately, the use of imagery can produce increases in errors and increased use of inferences in eyewitnesses' spoken reports (Bekerian & Dennett, 1997; for a discussion, see Stevenage & Memon, 1997).

Alternative retrieval cues can be used to access an event. For example, witnesses can be asked to recall an event in different temporal order or from different perspectives. Some researchers have found that witnesses can recall additional details if they deviate from the event script and describe the event from the end or the middle or if they describe its most memorable aspect (Fisher & Geiselman, 1992; Geiselman & Callot, 1990). However, in other studies, no additional details have surfaced when the witness recalls the event for a second time, in a different order (Memon, Wark, Bull, et al., 1997). One of the most controversial components of the original CI was that witnesses were asked to "recall" an event from the perspective of another witness or from another location at the scene. The instruction to change perspective typically does not yield additional details and can increase errors, particularly if witnesses do not understand what the interviewer wants them to do (Boon & Noon, 1994; Memon, Cronin, Eaves, & Bull, 1993). Fisher, Brennan, and McCauley (2002) suggest that changing perspectives could be potentially valuable for highly traumatized witnesses who might find it too stressful to describe the event from their own perspective. However, forensic investigators are uncomfortable with the instruction to change perspective, presumably because it could invite witnesses to speculate (Kebbell, Milne, & Wagstaff, 2001).

Evaluation of the CI

The CI has been examined in approximately 65 studies to date. A meta-analysis of 53 studies found a median increase of 34% in the amount of correct information generated in the CI as compared with a different interview model (Köhnken, Milne, Memon, & Bull, 1999). There was also an increase in incorrect details; we will return to this later. With the exception of two field studies, all the studies have tested volunteer witnesses (typically college students) in the laboratory. Witnesses observe either a live event or a videotape of a simulated crime. After a short delay (typically hours or days), the witnesses participate in a face-to-face interview. The witnesses receive either the CI or a control interview. The control is either a standard police interview or a structured interview that incorporates the phased approach referred to earlier. The interviews are tape recorded, transcribed, and then scored for the number of correct and incorrect statements. The accuracy of the reported statements is high and comparable for both types of interview.

Günter Köhnken and his colleagues in Germany (Köhnken, Schimmossek, Aschermann, & Höfer, 1995; Köhnken, Thurer, & Zorberbier, 1994) were the first to demonstrate the superiority of the CI over the structured interview. In their studies, the structured-interview group received training in basic communication skills that was comparable in quality and length to the CI group's training. The training included instruction on rapport building and the use of various types of questioning. In the 1994 study, both interviewees and interviewers were non-psychology students with no prior experience in investigative interviewing. The to-be-remembered event was a videotape showing a blood donation. Participants were tested a week after viewing the videotape. Each interviewer conducted one interview ($n = 30$). The CI significantly increased the amount of correctly recalled information over the structured interview without increasing the number of errors and confabulated (made-up) details. In a subsequent study with adult participants, a small increase in confabulated details was also noted (Köhnken et al., 1995).

Memon and colleagues (Memon, Wark, Holley, Bull, & Köhnken, 1997) directly examined whether the CI advantage was due to the use of the communication components of the revised CI (rapport building, transfer of control, and elements of the structured interview) or of the cognitive components (context reinstatement, imagery, reverse order, and reporting in detail). As in the Köhnken research, cognitive and structured interviewers received similar training, and each group was led to believe it was using the superior interview technique. A third group of interviewers served as the control and was not trained. Both trained groups elicited more correct information than the untrained group did. However, this was offset by the fact that they also produced a significantly higher number of errors and confabulations than the untrained group. These findings are important in themselves but also raise the question of what is an appropriate control group. The cognitive interviews produce more correct details than do interviews conducted by an untrained group of interviewers. However, a structured interview with some of the communication components of the CI built in can also yield increases in correct recall. The increase in errors that occasionally occurs could be somewhat problematic (for a discussion, see Memon & Stevenage, 1996; Memon, 2006). Some have argued that the production of incorrect as well as correct information suggests that the CI may be affecting report criteria (Memon & Higham, 1999; Roberts & Higham, 2002). Others argue that there is no suggestion that witnesses should lower their output criteria to produce unsure responses and interviewers should instruct witnesses not to guess or fabricate details (Fisher et al., 2002). It is important to note that accuracy rates typically do not differ between the CI and comparison groups.

The efficacy of the CI with nonstandard populations—notably, young children, the elderly, and people who are intellectually impaired—has also been examined. Given that the primary aim of the CI is to increase the amount of information retrieved, it may be the most effective procedure to use with young children, because children tend not to provide as much information as adults do. The results are somewhat mixed. The CI has been found to increase the amount of correct information recalled by children aged 7 to 11 years when the comparison group was a standard (untrained) group (Saywitz, Geiselman, & Bornstein, 1992). When the comparison is a structured interview, the CI increases correct information but can also increase errors in 8- to 9-year-olds (Memon, Wark, Bull, et al., 1997; Milne, Bull, Memon, & Köhnken, 1995).

More recently, Akehurst, Milne, and Köhnken (2003) examined whether the revised CI would aid the recall of children aged 8 to 9 years and 11 to 12 years after a 6-day delay. Children viewed a video of a shoplifting and were interviewed 4 hours or 6 days later. The CI led to an increase in correct recall as compared with a structured interview, with no increase in errors. There were no interactions involving age group or delay. As to the suitability of the CI for younger children, Holliday (2003a) reported that a modified version of the revised CI could increase the amount of correct information recalled in the narrative (free-recall) phase of the interview in 4- to 5-year-olds as compared with a structured interview. In a later study with 4- and 8-year-olds, Holliday (2003b) found that a CI given after postevent misinformation reduced children's reporting of misinformation in the interview (for a similar finding with 8- to 9-year-olds, see Memon, Wark, Bull, et al., 1997).

There is some evidence to suggest the CI can aid the recall of adults (Milne, Clare, & Bull, 1999) and children (Milne & Bull, 2006) with mild learning disabilities, although further research is required with this population, using larger sample sizes and people with a broader range of learning difficulties.

To date, there have only been two published studies of the efficacy of the CI when the witnesses are older adults. Mello and Fisher (1996) found the CI led to similar increases in correct recall when the participants were older adults (mean age = 72 years) but Searcy, Bartlett, Swanson, and Memon (2001) found no differences in correct identification (recognition) of a target when witnesses aged 62–79 years were interviewed using a procedure resembling the CI.

The failure to find an effect of CI on recognition (in this case, identification of a target) is consistent with earlier studies. In four separate studies, Fisher, Quigley, Brock, Chin, and Cutler (1990) found no advantage of the CI in recognition, but it did elicit better descriptions of the target as compared with a no-instruction control. Gwyer and Clifford (1997) compared the revised version of the CI with a structured interview and again found no reliable effects on recognition performance in target-present conditions but a reduction in false identifications in target-absent conditions in their short (48-hour) delay group (cf. Yarmey, 2004). This finding did not generalize to the long (96-hour) delay group.

These findings come as no surprise. The literature indicates that environmental manipulations of context are not effective in a recognition test when alternative cues are available. According to Smith and Vela (2001), this is because the influence of contextual cues will be reduced or will be outshone when there are strong retrieval cues present at the time of the memory test. This is referred to as the outshining hypothesis. For instance, in a recognition test in which a copy of the item to be remembered is provided, this item serves as a retrieval cue, and contextual cues are rendered ineffective. When the task is to recall an item of information in the absence of a specific retrieval cue, the reinstatement of context should guide memory (Smith, 1994). However, as pointed out by Fisher and Schreiber (in press), the outshining hypothesis leads to the prediction that experimental manipulations should have smaller effects in target-present than in target-absent conditions.

Future studies should examine whether witnesses interviewed with the revised CI are more likely to make correct rejections and whether the effect of a CI in an identification situation will vary as a function of retention interval (Gwyer & Clifford, 1997) and other relevant system and estimator variables.

Application/Training

Police officers complain that eyewitnesses seldom provide sufficient information (Kebbell & Milne, 1998). The CI has proved to be a prime system variable in that a full and accurate eyewitness statement may determine whether or not a case is solved. The question is, what impact has the CI had on interviewing practice?

Despite the extensive scientific research on the CI, knowledge and application of it is not widespread among investigators in the United States, and it does not appear to have had a substantial impact on the methods police officers use to interview witnesses (Fisher & Schreiber, in press). Nevertheless, personnel from police and nonpolice organizations have received training in the technique. These organizations include the FBI; the National Transportation Safety Board; the Department of Homeland Security; the Federal Department of Law Enforcement; and the Bureau of Alcohol, Tobacco, and Firearms. NASA personnel will receive such training in the near future. The training has varied across states and differs between federal and state training academies. Fisher and Schreiber (in press) note that federal investigators receive 18 hours of training in interviewing, including techniques for interrogating suspects and nonpsychological topics such as the legal aspects of interviewing.

In England and Wales, the CI was introduced in a booklet to every police officer as part of the national investigative interviewing package in 1992. However, while Britain has some good examples of police training in the CI, with input to the trainers from researchers, the training is typically limited to the detective ranks or is only provided in a minimal, introductory form to junior officers (see Milne & Bull, 2006). A survey of police officers (Kebbell, Milne, & Wagstaff, 1999) suggested that relatively few officers used the full CI in practice. Training programs have also been developed in other European countries, as well as in Australia, New Zealand, and Israel (Fisher, 2005). The efficacy of the CI has also recently been demonstrated in Brazil (Stein & Memon, in press), with the aim of introducing it to the Brazilian police and judiciary in the near future.

Given the extensive research on the CI and the light it has shed on faulty interviewing practices, have police interviews improved in the 20 years since the CI was first introduced? In a recent analysis of police interview techniques, Fisher and Schreiber (in press) asked 23 Miami detectives experienced in investigations of robbery, sexual assault, homicide, and internal affairs to tape record their witness interviews. Analysis of these interviews revealed techniques and behaviors similar to those identified 20 years earlier. This was particularly disappointing in view of the scientific progress made in the field and the efforts by Fisher and Geiselman to disseminate their findings to practitioners and to implement training programs.

The picture is just as bleak across the Atlantic. Clarke and Milne (2001) conducted a national evaluation of investigative interviewing training (the Planning, Engage, Account, Closure, Evaluation—or PEACE—model) in England and Wales to see if it had improved workplace practice. The PEACE model provides two ways of obtaining an interviewee's account: the CI and conversation management. The latter involves asking witnesses to give their own account of events; the interviewer then selects specific topics from the account and questions the witnesses in a logical sequence. Clarke and Milne (2001) found little evidence of any cognitive interviewing taking place. Most officers seem preoccupied with getting a statement from the witness and asking closed questions. One reason for the lack of development in witness-interviewing skills is that resources have targeted the use of interrogative techniques or suspect interviews at the expense of gathering information from cooperative witnesses (Milne & Bull, 2006).

Resources need to be directed toward training in witness-interviewing practices. Milne and Bull (2006) argue that this will involve procedural changes in collecting evidence in the United Kingdom, such as electronic recording of all witness interviews to maintain an accurate record of the original account, assessment of training and supervision of witness, and victim interviews to ensure that appropriate techniques are used. With respect to the United States, R.P. Fisher (personal communication, March 28, 2006) has noted that nonpolice groups, such as engineers, have displayed a willingness to use CI in investigations, suggesting that perhaps those with an academic background or a motivation to use investigative techniques to arrive at solutions find it easier to understand the theory behind the CI. Following this line of reasoning, perhaps police officers with specialist skills (homicide, child protection) might benefit more from training in the CI. However, those who are specialists may already have an established protocol for interviewing and thus be less willing to adopt new techniques (Memon, Milne, Holley, Bull, & Köhnken, 1994).

We advocate a two-tiered approach to training. First, there is a need for more extensive training programs on witness-interviewing techniques for new police officers. Training and examples of how faulty witness testimony can contribute to miscarriages of justice might also prove fruitful (see Savage & Milne, in press). The monitoring and assessment of witness interviews (e.g., recording) is essential. A second approach is to present trainees with a simpler, more accessible version of the cognitive interview (e.g., Davis, McMahon, & Greenwood, 2004) to encourage wider use.

Identifying Criminal Suspects

The identification of a criminal suspect can be the most important eyewitness evidence that is presented at a trial. This is especially true when the eyewitness claims to have seen the suspect commit the criminal act. In that case, the eyewitness-identification testimony is direct evidence of guilt in the sense that the accuracy of the identification has a one-to-one relationship to the ultimate issue of whether the suspect committed the crime. In other situations, eyewitness identification evidence may be circumstantial—for instance, if the eyewitness only saw the person in the vicinity of the crime or saw the person leaving a building at a certain time. In these cases, other types of evidence are needed to complete the inference that the person who was seen is the same person as the one who committed the crime. Regardless of whether the identification is direct or circumstantial, those who observe identification testimony (for example, jurors) are likely to accept it as accurate if the eyewitness is confident and consistent (e.g., Berman & Cutler, 1996; Bradfield & Wells, 2000; Brigham & Bothwell, 1983; Cutler, Penrod, & Stuve, 1988; Lindsay, Lim, Mirando, & Cully, 1986; Lindsay, Wells, & O'Connor, 1989; Lindsay, Wells, & Rumpel, 1981; Maas, Brigham, & West, 1985; Wells & Leippe, 1981; Wells, Lindsay, & Ferguson, 1979).

Lineups

A primary method for obtaining identifications of criminal suspects is the use of the lineup. Lineups can be either live, as commonly seen on TV shows, or photographic. In the experience of the first and third authors, most lineups in the United States are conducted using photographs. At its simplest level, a lineup involves placing a suspect among distracters (called fillers) and asking the eyewitness if he or she can identify the target. The lineup is more complex than it at first appears. Understanding how mistaken identifications can occur with lineups and what kinds of system improvements can be made to prevent mistakes requires an understanding of the structural properties of lineups and their possible outcome distributions.

Lineup Structure. Regardless of whether there is more than one culprit, or target, a lineup should contain only one suspect, with the remaining members being known-innocent fillers (Wells & Turtle, 1986). It is critical to keep in mind that the suspect might or might not be the target (i.e., might or might not be the actual culprit). Hence, we will refer to two possible states of truth: (a) the suspect is the target, and (b) the suspect is not the target. Because there is only one suspect per lineup, these two states of truth are equivalent to target-present and target-absent lineups. In a target-present lineup, two kinds of errors can be made: (a) an incorrect rejection (making no identification), and (b) the identification of a filler. Note that one cannot mistakenly identify an innocent suspect in a target-present lineup. The only time an eyewitness can mistakenly identify an innocent suspect is in a target-absent lineup. Target-absent lineups can also result in filler-identification errors, but these errors would not result in charges being brought against an innocent person. We reserve the term "mistaken identification" to refer to the identification of an innocent suspect; the identification of anyone other than the suspect is called filler identification. Thus, the structural properties of a lineup produce the set of possible outcomes shown in Table 3. In an experiment, participant witnesses are shown either a target-present or target-absent lineup to simulate the real-world fact of an unknown probability that the police are focusing on an innocent suspect. The proportion of target-present and target-absent lineups (the target-present base rate) is commonly 50/50 for experiments, but Bayesian statistics permit quantitative analyses of what happens across all possible base rates for any given experiment (see Wells & Lindsay, 1980; Wells & Olson, 2002; Wells & Turtle, 1986).

Typical Outcome Distributions. As would be expected from better-than-chance performance, experiments typically show that accurate identifications are more likely than inaccurate identifications and that true rejections are more likely than are false rejections (Clark, 2003; Wells & Lindsay, 1980; Wells & Olson, 2002). Notice, however, that there are two types of filler identifications. Filler identification Type 2 is a "miss" in the sense that the target was present and could have been chosen but the eyewitness picked someone else. Filler identification Type 1 is an accurate rejection in the sense that the suspect is innocent and the eyewitness did not pick him or her. In general, experiments show that Type 2 filler identifications are more likely than

TABLE 3
Possible Outcomes From a Lineup

	Response of Eyewitness		
State of truth	Identification of suspect	Identification of filler	No identification
Suspect not target	Accurate identification	Filler identification type 1	False rejection
Suspect is target	Mistaken identification	Filler identification type 2	True rejection

Note. Filler identification type 1 can be construed as an accurate rejection in the sense that the target was not present and the eyewitness did not pick him or her. Filler identification type 2 is a ''miss'' in the sense that the target was present but was not picked. Source: Charman & Wells (2006).

are Type 1 filler identifications (Wells & Olson, 2002). This makes sense and fits well with the concept of relative judgments (Wells, 1984), in which it is presumed that eyewitnesses tend to select the person who looks most like the target. When the target is absent, the chances increase that one of the fillers will be perceived as looking like the target. Usually, eyewitness-identification performance is calculated by the extent to which accurate identifications exceed mistaken identifications and true rejections exceed false rejections. However, the rate of mistaken identifications can be decreased without increasing correct rejections by shifting identifications to fillers in the target-absent lineup. This is a key to understanding how careful selection of fillers for lineups can reduce mistaken identifications even if it does not reduce the propensity of eyewitnesses to attempt identifications from target-absent lineups.

Target Removal Without Replacement. The relative-judgment conceptualization (Wells, 1984) has permeated the literature on lineups. It simply states that eyewitnesses have a tendency to identify a person from the lineup who looks most like their memory of the target relative to the other lineup members. As long as the actual target is in the lineup, the relative-judgment process should work well. However, if the actual target is not in the lineup, problems ensue, because there will always be someone who looks more like the target than the other lineup members. Various results have been interpreted as supporting the relative-judgment conceptualization, but the removal-without-replacement (RWR) effect is the best evidence in support of the relative-judgment conceptualization.

In the original demonstration of the RWR effect (Wells, 1993), eyewitnesses viewed either a six-person lineup that contained the target or a five-person lineup in which the target was removed and not replaced with anyone. In both conditions, the eyewitnesses were instructed that the target might not be present (see following section on pre-lineup instructions). When the target was present, 54% picked the target, 21% selected no one, and 25% selected fillers. Wells reasoned that if the 54% represented true recognition rather than a relative judgment, removal of the target should result in the 54% joining the 21% in picking no one. When the target was removed, however, only 32% selected no one, and 68% selected fillers. Thus, among the 54% selecting the target when the target was present, it is estimated that 79.6% of them (43%/54%) would have selected one of the fillers in the absence of the target. Recent data show the RWR effect to be robust across a variety of conditions, and the magnitude of the effect appears to be greater when memory is weaker (Clark & Davey, 2005; MacLin, Wells, & Phelan, 2004). There remains some debate about the psychological processes underlying the RWR effect. Ebbessen and Flowe (n.d.), for instance, suggest that it could simply represent a downward criterion shift that occurs when the target is removed. Regardless of the interpretation, the RWR effect illustrates the substantial risk that accrues to an innocent suspect when the actual target is not present.

The effect also further illuminates the problem of filler selections that we noted earlier in the discussion of archival studies using police files, in which one third of positive identifications by witnesses were identifications of innocent fillers. In the American archival study, Behrman and Davey (2001) found that nearly a quarter of witnesses selected a filler (and 50% selected the suspect). Thus, the average filler was selected by 5% of witnesses—what might be termed "bad guesses" (at least in the sense that witnesses' memories were not good enough to avoid errors; Penrod, 2003). Of course, in a perfectly fair array, one would have to assume that at least another 5% of witnesses would "guess" the suspect. These selections might be characterized in various ways: Steblay et al. (2001) called them "calculated guesses" and Penrod called them "lucky guesses."

As we discuss later, there are reasons to believe that many lineups are not fair and that calculated/lucky guesses produce many suspect identifications that look like "hits" but are really the product of biased arrays and witness guessing. Steblay et al. (2001) reported, for instance, that in studies of target-absent simultaneous arrays in which a filler similar to the suspect was designated the "innocent suspect," that person was picked by 27% of witnesses (across all studies, one of the six fillers—including the suspect—was picked by 51% of witnesses). One might expect that in a fair lineup the innocent filler would be selected by 8.5% (51%/6) of witnesses instead of 27%. The much higher rate of suspect identification suggests that the witnesses had some memory for the appearance of the missing target but not enough of a memory to avoid mistakenly identifying an innocent person.

Pre-Lineup Instructions. One of the first and most fundamental lineup system variables to be tested empirically was the instruction (or warning) to eyewitnesses that the target might or might not be in the lineup. Malpass and Devine (1981) used both target-absent lineups and target-present lineups; they either gave the pre-lineup instruction that the target might or might not be present or gave no instruction. When participants viewed a target-present lineup, the instruction had little effect on the distribution of responses. When they viewed a target-absent lineup, however, the instruction reduced choosing rates dramatically. This general pattern, in which the instruction reduces the chances of both mistaken identifications and filler identifications, has been replicated extensively (see meta-analysis by Steblay, 1997). A more recent meta-analysis indicates that accurate identification rates in target-present lineups might be slightly harmed by the instruction, but the decline in accurate identifications when the target is present is much smaller than the decline in mistaken identifications when the target is absent (Clark, 2005).

Selection of Fillers. The characteristics of the fillers used in a lineup have a strong influence on the chances that an innocent suspect will be identified in a target-absent lineup. In general, if the innocent suspect fits the description of the target and the fillers do not, the innocent suspect is likely to be mistakenly identified. The first empirical demonstration by Lindsay and Wells (1980) was followed by a debate about the optimal criteria for selecting fillers. Two primary strategies for selecting fillers have been advocated. One is to select fillers who resemble the suspect. Luus and Wells (1991) argued against this strategy because it has no "stopping point" and also because it risks creating a lineup of clones, which would reduce accurate identification rates for target-present lineups. Wells, Rydell, and Seelau (1993) found that selecting fillers on the basis of the description given by the eyewitnesses managed to protect the innocent suspect in target-absent lineups without harming accurate identification rates in target-present lineups. On the other hand, selecting fillers on the basis of their resemblance to the suspect harmed hit rates with no additional protection for the innocent suspect in target-absent lineups.

Wogalter, Marwitz, and Leonard (1992) presented another argument against selecting fillers on the basis of their resemblance to the suspect: The "backfire effect" refers to the idea that, somewhat ironically, the suspect might stand out if he or she was the basis for selecting the fillers in the lineup, because the suspect represents the central tendency or origin of the lineup. Clark and Tunnicliff (2001) reported evidence for the backfire effect. However, eyewitnesses' descriptions of the target are often sparse and sometimes do not even match the characteristics of the suspect (Lindsay, Martin, & Webber, 1994; Meissner, Sporer, & Schooler, in press; Sporer, 1996, in press). The general recommendation for selecting fillers for lineups has been to use the eyewitness's description of the target and to take any additional measures needed to make sure that the suspect does not stand out in the lineup (Wells et al., 1998).

Along with these strategies for selecting fillers, various techniques to assess lineup fairness by using "mock witnesses" have been developed. The task of a mock witness is to examine the lineup and try to discern which person is the suspect. From this mock-witness paradigm, various metrics have been developed to assess the extent to which the suspect stands out unfairly (Malpass & Lindsay, 1999). In lab studies, the mock-witness paradigm appears to be sensitive to lineup bias and is relatively robust across variations in lineup procedure (e.g., simultaneous vs. sequential procedures; see McQuiston & Malpass, 2002). Studies of photo arrays and lineups from actual cases using the mock-witness method reveal that arrays are frequently biased against suspects, who are picked more than twice as often (relative to the fillers) as one would expect by chance alone (Brigham et al., 1999; Valentine & Heaton, 1999; Wells & Bradfield, 1999b).

Lineup Size. A common practice in the United States is to use five or six persons (a suspect plus four or five fillers) in a live lineup and six or eight photos in a photo lineup. For purposes of this discussion of lineup size, we will assume that each lineup member is viable in the sense that the fillers are selected to fit the description and in other ways do not make the suspect stand out. Given a set of properly selected lineup fillers, there is no reason to believe that an innocent suspect has a greater chance than any of the fillers to be identified by an eyewitness. Hence, eyewitness researchers have adopted the assumption that the chances of a mistaken identification are $(1/N) \times p(\mathrm{I})$, where N is the number of lineup members and $p(\mathrm{I})$ is the probability that an eyewitness will make an identification (see Doob & Kirshenbaum, 1973; Wells, Leippe, & Ostrom, 1979). Note that increasing lineup size reduces the chances of a mistaken identification in a negatively decelerating fashion (i.e., each additional lineup member reduces the chances of a mistaken identification less than the previous addition did). Because of this negative deceleration, the addition of persons to the lineup brings diminishing returns. Thus, adding six additional members to a six-person lineup reduces the chances of mistaken identification from 16.7% to 8.3% (i.e., among those making an identification). But, adding six members to a 12-person lineup reduces the chances of mistaken identification from 8.3% to 5.5%.

If reducing the chances of a mistaken identification were the only consideration, increasing the size of the lineup to a very high number is a good idea, even with diminishing returns. But the formula speaks only to mistaken identifications from target-absent lineups and not to the chances of accurate identifications from target-present lineups. The idea of a system variable improvement for lineups is to reduce the chances of a mistaken identification without harming the chances of an accurate identification. Thus, the critical question is what happens to accurate identifications as a function of increasing lineup size? The eyewitness-identification literature has not derived a precise function relating lineup size to accurate-identification rates. Levi (2002) reported no drop in accurate-identification rates when lineup sizes were increased from 10 to 40 persons. In fact, the literature includes reports of eyewitnesses being able to view up to about 300 photos with little reduction in the chances of an accurate identification (Ellis, Shepherd, Flin, Shepherd, & Davies, 1989; Lindsay, Nosworthy, Martin & Martynuck, 1994). These results are consistent with the general observation that identifications of the target from target-present lineups are not as sensitive to lineup variations as mistaken identifications from target-absent lineups are (Charman & Wells, 2006). For example, the "might or might not be present" instructions have little effect on accurate identifications from target-present lineups but appreciably reduce identifications from target-absent lineups (Steblay, 1997). Similarly, the use of a filler-biased lineup has little effect on accurate identifications from target-present lineups but increases mistaken identifications from target-absent lineups (Wells, Rydell, & Seelau, 1993). Also, suggestive influences from lineup administrators appear to have little effect when eyewitnesses view a target-present lineup but have a strong effect when the eyewitnesses view a target-absent lineup (Haw & Fisher, 2004). More systematic research is needed before it will be possible to conclude that lineup sizes can easily be raised to 20 or more persons without harming accurate identification rates, but there appears to be great promise in the simple idea of increasing the nominal size of lineups.

Double-Blind Lineups. Police conducting a lineup has been likened to psychologists conducting an experiment (Wells & Luus, 1990). One element of this rich analogy is the idea of the double-blind lineup (Wells, 1988). Normally, a lineup is conducted by the case detective, who also assembled the lineup and knows which person is the suspect and which people are merely fillers. The psychological literature on experimenter-expectancy effects reveals the dangers of permitting a person who knows the correct, desired, or expected answer to administer a face-to-face test (Harris & Rosenthal, 1985), and yet this is standard practice for lineups. Experiments have shown that when the lineup administrator is led to believe that a particular lineup member (randomly selected) is the suspect, the chances that the eyewitness will identify that person are increased (Haw & Fisher, 2004; Phillips, McAuliff, Kovera, & Cutler, 1999). Furthermore, when the eyewitness selects the person whom the lineup administrator was led to believe is the target, the eyewitness expresses higher levels of confidence in the identification (Garrioch & Brimacombe, 2001).

The idea of the double-blind lineup is straightforward: The person who administers the lineup should not be aware of which lineup member is the suspect and which members are fillers (Wells et al., 1998). This recommendation does not presume any intention or awareness on the part of the lineup administrator to influence the eyewitness. Some police jurisdictions might be concerned about manpower issues involved in using an independent lineup administrator. Because most lineups in the United States are actually photo spreads of some sort rather than live lineups, an alternative to using a double-blind administrator is to have a laptop computer administer the lineup, thereby effectively eliminating any possible influence from the lineup administrator (for a description of such a program, see MacLin, Zimmerman, & Malpass, 2005).

Sequential Lineups. An alternative to the traditional police lineup, the sequential lineup, was introduced in the mid-1980s (Lindsay & Wells, 1985). Unlike the traditional police lineup, in which all members are presented to the eyewitness at once (simultaneous lineup), the sequential lineup presents the lineup members to the eyewitness one at a time. The eyewitness is told that he or she will view a number of people—the number is not specified. The witness makes a decision on each lineup member (yes, no, or not sure) before the next lineup member appears. The theory behind the sequential lineup is that it prevents eyewitnesses from relying on relative judgments, in which one lineup member is compared with another and the one most similar to the target is picked. Although the eyewitness can compare the lineup member currently being viewed with those already seen, there is a chance that a lineup member yet to come might look even more similar to the target. The initial results indicated support for a sequential-superiority effect in which identifications from target-absent lineups diminished while identifications of the target from target-present lineups remained largely the same.

Years of additional experiments culminated in a meta-analysis that aggregated data across 4,145 participant witnesses (Steblay et al., 2001). The meta-analysis supported the original observation of lower mistaken identification rates for the sequential than for the simultaneous lineup; however, there was also a reduction in accurate identifications of the target from the target-present lineups. In general, the sequential procedure appears to result in fewer identification attempts overall compared with the simultaneous procedure. Although the sequential procedure reduced mistaken identifications at a greater rate than it did accurate identifications, this shift in performance is consistent with a criterion shift in which eyewitnesses set a higher criterion for identification with the sequential than with the simultaneous procedure (Meissner, Tredoux, Parker, & MacLin, 2005). However, these results are also consistent with a shift away from relative judgments.

Recall that the RWR effect indicates that some proportion of accurate identifications appears to result from relative judgments rather than true recognition. Thus, a shift away from relative judgments is likely to result in fewer accurate identifications as well as fewer mistaken identifications. An argument can be made that a more conservative lineup test (whether owing to a higher criterion for making a positive identification or to a reduced reliance on relative judgments) is desirable, as mistaken identification is the primary cause of convictions of the innocent. The trade-off of accurate and mistaken identifications ultimately is a decision for policymakers, not scientists. However, Steblay et al. (2001) and Penrod (2003) argue that any losses of accurate identifications that result from reduced reliance on relative judgments are merely reductions in lucky or calculated guesses.

A recent field experiment involving actual lineups conducted in Illinois (Illinois Pilot Program, 2006) has been touted as a comparison of the sequential lineup and the traditional simultaneous lineup. The authors' report on the experiment interprets its results as indicating that the traditional simultaneous lineups yielded fewer filler identifications and more suspect identifications than did the sequential lineups. In fact, however, this two-condition experiment actually confounded several variables. Perhaps the most important confound was that the simultaneous lineups were never conducted using double-blind procedures but were always conducted by the case detectives. The sequential lineups, in contrast, were always conducted using the double-blind method. Thus, the low filler rate obtained in the simultaneous lineups could have been the result of not using double-blind procedures. Consistent with this concern, it should be noted that the double-blind sequential-lineup data in the Illinois Pilot Program conformed quite well with data obtained using the double-blind sequential procedure in the Hennepin County (Minnesota) pilot project (about 8% filler identification rates; see Klobuchar, Steblay, & Caligiuri, in press). In contrast, the very low filler rate reported in the Illinois Pilot Program using the nonblind simultaneous procedure (about 3%) is an extreme outlier from the approximately 20% rate found in other jurisdictions with simultaneous lineups (see Behrman & Davey, 2001; Slater, 1994; Valentine et al., 2003; Wright & McDaid, 1996). The profoundly low filler-identification rate for simultaneous lineups reported in the Illinois Pilot Program suggests a suppression of filler identifications and/or a reluctance to report filler identifications by the nonblind lineup administrators. Thus, we are reluctant at this time to consider the Illinois Pilot Program to be an interpretable test of the simultaneous versus sequential procedure.

Composites

When there is no clear suspect, investigators sometimes resort to the use of sketch artists or composite faces. Little systematic research on sketch artists exists, in part because variance across artists (e.g., in their abilities) is presumed to be significant and a fairly large sample would be required to reach generalizable conclusions. Considerable research exists, however, on composite production systems, which are increasingly being used by crime investigators in place of sketch artists. Composite production systems create faces by selecting features (e.g., nose, eyes, chin, head shape, hair, mouth, brows, ears) and combining them into a face. One of the original systems, Identi-Kit, used line drawings of facial features on transparencies. An accompanying booklet displayed all the possible features, and the eyewitness selected features that were then overlaid on each other to form a complete face. A later system, Photo-Fit, used the same system, except that the features were black-and-white photos of actual facial features instead of line drawings.

In recent years, computer software programs have replaced transparency-based composite systems. Examples of such software are E-Fit, Evo-FIT, CD-Fit, and Mac a Mug (Frowd et al., 2005). The FACES program is currently popular among U.S. law enforcement agencies (Cote, 1998). FACES includes 361 hair selections, 63 head shapes, 42 forehead lines, 410 sets of eyebrows, 514 sets of eyes, 593 noses, 561 sets of lips, 416 jaw shapes, 145 moustaches, 152 beards, 33 goatees, 127 sets of eyeglasses, 70 eye lines, 147 smile lines, 50 mouth lines, and 40 chin lines. In each feature category, a selection button permits the user to view subsets of the feature that meet a particular description. For instance, eyes are divided into the subsets narrow, deep set, overhanging lids, heavy lids, average blue or green, almond-shaped blue or green, average brown, almond-shaped brown, and bulging. Noses are divided into the subsets of narrow, average with round base, average with broad base, average pointed, hooked nostrils not showing, hooked nostrils showing, slightly flared nostrils, very flared nostrils, round (bulbous), average large, wide base with nostrils showing, and wide base with nostrils not showing. In addition, controls permit the features to be moved up or down and closer or farther apart, and to be made larger or smaller. The features are displayed on one side of the computer screen, and the face is built on the other side. When a feature is clicked, it appears on the face. To make changes—for example, in the eyes—one simply clicks a different set of eyes, and those on the face are replaced with the new ones.

All composite systems use a part-to-whole method to build the face: The eyewitness constructs a face by selecting features and assembling them. Numerous face-recognition researchers have noted that this method may conflict with the natural way faces are encoded in memory—namely, in a holistic manner (e.g., Tanaka & Farah, 1993; Tanaka & Sengco, 1997; Wells & Hryciw, 1984). Research experiments generally indicate that composite faces tend to be rather poor likenesses of the original faces (e.g., Bruce, Ness, Hancock, Newman, & Rarity, 2002; Ellis, Davies, & Shephard, 1978; Kovera, Penrod, Pappas, & Thill, 1997). The research by Kovera et al. illustrates the difficulty of generating a composite that resembles the intended target. The researchers used a set of 50 composite images of the faces of high-school classmates and faculty created by former students. Other graduates of the same schools judged the composites' quality. The judges were told that some of the composites were of former high-school classmates; they were asked to identify them, rating the faces' familiarity and their own confidence in that assessment and, where possible, giving names. Ratings of familiarity and confidence did not differentiate significantly between the known and unknown faces, and only 3 of the 167 names offered for the composites proved to be correct! Ratings by the composite constructors of their familiarity with the targets and their assessments of the quality of their composites were unrelated to identification accuracy on any measure. The researchers concluded that "the findings . . . raise doubts about the likelihood that composites prepared under field conditions will yield a pinpointed identification of a perpetrator by individuals who know the perpetrator" (Kovera et al., 1997, p. 245).

Although early research using the Identi-Kit and Photo-Fit suggested that the poor likenesses might be due to the composite systems themselves (e.g., too few choices of features; Ellis et al., 1978), there is an emerging consensus that people simply do not have good memories for isolated facial features and that any system that requires parts-to-whole-face recall will be severely limited. Furthermore, research suggests that having eyewitnesses build a composite face can damage memory for the original face and make the witnesses less able to recognize the original target face in a later lineup (Wells, Charman, & Olson, 2005). Similar effects have been observed for giving verbal descriptions of faces, a phenomenon called the *verbal overshadowing effect* (originally demonstrated by Schooler & Engstler-Schooler, 1990; and see meta-analysis by Meissner & Brigham, 2001b).

Recent research has produced some encouraging results for cases in which multiple eyewitnesses independently produce composites. In such cases, morphing the individual composites produces a new face that is a better likeness of the person than is any individual composite (Bruce et al., 2002; Hasel & Wells, in press). But even the morph of four individual composites does not produce a dramatic likeness of the original face. Hasel and Wells reported that the ability to pick the original target from sets of four alternative faces was 35% for individual composites and 48% for four-composite morphs (chance = 25%).

POSTDICTION VARIABLES

Postdiction variables are neither system nor estimator variables in the traditional sense, because they are not presumed to causally affect the accuracy of eyewitnesses. These variables are measurable products that correlate with the accuracy of eyewitnesses in a noncausal manner. The most researched of these is the confidence (certainty) of the eyewitness. Another postdiction variable is response latency—specifically, how long the eyewitness takes to make an identification. The third postdiction variable that we review here is self-reported decision process.

Confidence

The confidence an eyewitness expresses in his or her identification is one of the most researched questions in the study of eyewitnesses. First, there is a strong intuitive appeal to the idea that confidence and accuracy should be closely related. Second, courts have explicitly endorsed the idea that the reliability of an eyewitness should be gauged at least in part by the person's confidence, a tenet advocated by the U.S. Supreme Court (*Manson v. Braithwaite*, 1977). Third, even in the absence of instructions to pay attention to eyewitness confidence, participant jurors rely heavily on the confidence of the eyewitness in deciding whether he or she made an accurate identification (e.g., Bradfield & Wells, 2000; Fox & Walters, 1986; Lindsay et al., 1986; Lindsay et al., 1989; Lindsay et al., 1981; Wells, Ferguson, & Lindsay, 1981; Wells et al., 1979).

Initially, eyewitness researchers focused on the relationship between eyewitness-identification confidence and eyewitness-identification accuracy (Wells & Murray, 1984). This was a useful starting point, but it is now clear that the relationship between confidence and accuracy varies greatly as a function of many other factors. For instance, it depends, in part, on how similar the mistakenly identified person is to the actual target (Lindsay, 1986). The confidence–accuracy relationship is generally higher when memory strength is stronger rather than weaker (Deffenbacher, 1980); when it is calculated only among those who make an identification rather than among both those who make an identification and those who do not (Sporer, Penrod, Read, & Cutler, 1995); and when it is calculated across witnesses under different viewing conditions rather than among witnesses who had the same viewing conditions (Read, Vokey, & Hammersley, 1990).

In their meta-analysis of 30 studies involving a total of 4,036 participant witnesses, Sporer et al. (1995) estimated that the confidence–accuracy correlation among choosers could be as high as +.41. Wells, Olson, & Charman (2002) note that a .41 point-biserial correlation (a correlation between a two-level variable and a continuous variable) between confidence and accuracy in eyewitness identification is less than the point-biserial correlation between height and gender in humans. Nevertheless, under conditions of uncertainty, a postdiction variable that has a .41 correlation to a criterion variable is not something that should be ignored. In fact, the American Psychology-Law Society's white paper on lineups endorses the idea of making a clear record of the confidence of an eyewitness that triers-of-fact might later use (Wells et al., 1998).

Accuracy of Highly Confident Witnesses

Though confidence–accuracy correlations are sometimes relatively high, most research yields relatively low correlations. Attempts have been made to increase the correlation through accountability, context reinstatement, and other thought manipulations, but none has been successful, and such measures commonly have the reverse effect of harming the confidence–accuracy relationship (Robinson & Johnson, 1998). Some have argued that despite the generally weak confidence–accuracy correlation, accuracy may be very high among the most confident witnesses. One analytic method that addresses this question uses calibration methods that measure peoples' confidence on a percentage scale (zero, 10%, 20%, 30%, and so on) and then clumps people together at different levels of confidence to assess their accuracy (see Brewer, Keast, & Rishworth, 2002; Brewer, Weber, & Semmler, 2005; Brewer & Wells, 2006; Cutler & Penrod, 1989; Juslin, Olsson, & Winman, 1996; Weber & Brewer, 2003, 2004).

Cutler and Penrod found witness overconfidence of 10 to 20% (that is, witnesses were making 10%–20% more errors than their confidence levels indicated). Juslin et al. (1996) found that confidence scores were roughly comparable to accuracy scores; in particular, in a 95% confidence group, judgments were 85 to 90% accurate (the exact numbers are not reported—numbers are estimated from figures). Although these numbers look promising, even in the 95% confidence group there appear to be 10 to 15% errors; errors are much higher—with greater overconfidence—at lower confidence levels.

Other researchers have found less promising results. Though the published numbers are slightly ambiguous, it appears that the top 21% most confident witnesses in Brigham, Maas, Snyder, and Spaulding (1982) were 85% correct. Brewer et al. (2002) found that eyewitnesses who were very confident in the accuracy of their identifications (95% certain) were about 70 to 75% correct—that is, high error rates and substantial overconfidence. In a 1987 study by Fleet, Brigham, and Bothwell, 75% of subjects who rated themselves as extremely confident were accurate. Brigham (1990) found a 74% accuracy rate for the top 27% most confident witnesses. Bornstein and Zickafoose (1999) reported that they found overconfidence in both general-knowledge domains and eyewitness-memory domains and that the two were correlated. The latter finding suggests that confidence has an individual-difference component that can be independent of the task. Research by Perfect and Hollins (1996) suggests that poor confidence–accuracy relationships are at least partly attributable to people's lack of insight regarding their general abilities in the eyewitness domain.

The general point is that these results are consistent with other measures of the confidence–accuracy relationship. Even the calibration approach does not uniformly support the notion that confidence is a highly reliable indicator of accuracy. Error rates can be high among even the most confident witnesses. Furthermore, these numbers presume that the criminal justice system would skim off only the most confident witnesses and that none of those witnesses would have had their confidence artificially boosted.

The Problem Grows Worse

Imagine that prosecutors are skimming only the most confident witnesses; there is no artificial confidence-boosting among the witnesses; and we have reliable measures of confidence, not the vague verbal reports currently obtained by police. Among these highly confident witnesses, the results above indicate that 20 to 30% could be in error. But even if the error rate is only 10% for these highly selected and most confident witnesses, they will all appear highly confident to jurors—so confidence cannot help the jurors figure out which witnesses have made errors. Indeed, the simple correlation between confidence and accuracy for these witnesses will be much worse than among all witnesses, because there is very little variability in confidence and maybe no useful variance. Though it is tempting to conclude that jurors might be entitled to assume a fairly high base rate of accuracy among these highly confident witnesses (even if confidence cannot aid them in differentiating accurate and inaccurate witnesses), the pleading effect discussed earlier suggests that it would not be safe to conclude that the accuracy rate is fairly high; indeed, the accuracy rate could be fairly low, because the guilty defendants facing confident witnesses have already pleaded guilty. In short, the research results and logic call into question the notion that witness confidence can be of significant assistance to jurors.

Even if the research showed that eyewitness-identification confidence and accuracy are related at a level that could have practical utility, this conclusion would come with another huge caveat. Wells and Bradfield (1998) showed that giving confirming feedback to eyewitnesses who had made mistaken identifications (e.g., "Good, you identified the suspect") produces profound distortions in their retrospective judgments, including their recollections of how confident they were when they made their identification, how good a view they had when they witnessed the event, and how much attention they devoted to the target's face during the event.

The idea that eyewitness confidence can be driven by variables that are independent of accuracy has theoretical roots in Leippe's (1980) early analysis of the problem, but the fact that other testimony-relevant variables (such as self-reports of attention and view) are also malleable is a startling revelation. There are numerous replications of this phenomenon, known as the *post-identification feedback effect* (Bradfield, Wells, & Olson, 2002; Dixon & Memon, 2005; Hafstad, Memon, & Logie, 2004; Neuschatz et al., 2005; Semmler, Brewer, & Wells, 2004; Wells & Bradfield, 1998, 1999a; Wells, Olson, & Charman, 2003; also see a meta-analysis by Douglass and Steblay, in press). The post-identification feedback effect occurs even if the feedback is delayed for 48 hours (Wells et al., 2003). The effect occurs for both positive identifications and "not there" decisions (Semmler et al., 2004), and the effect occurs for both the elderly (Neuschatz et al., 2005) and young children (Hafsted et al., 2004). Importantly, the confidence-inflating effect of confirming feedback is greater for eyewitnesses who have made a mistaken identification than for those who have made an accurate identification; as a result, confirmatory post-identification feedback harms the accuracy–confidence relationship (Bradfield et al., 2002). Furthermore, a recent experiment showed that the post-identification feedback effect occurs for actual eyewitnesses to real crimes (Wright & Skagerberg, in press).

The post-identification feedback effect is of considerable practical import, because it is a common practice for lineup administrators to give eyewitnesses feedback about their identifications. When an eyewitness has received some form of feedback before being asked about his or her confidence in the identification, the confidence statement is contaminated. Eyewitnesses tend to believe that the feedback did not affect them; however, those who report that the feedback did not affect their response to the retrospective confidence question are nevertheless affected just as much as are the smaller portion of witnesses who report that it might have affected them (Wells & Bradfield, 1998). Fortunately, if the eyewitness is asked to indicate his or her confidence level before receiving feedback, this tends to inoculate the eyewitness against post-identification feedback effects (Wells & Bradfield, 1999a). The need for immediate measures of confidence is further indicated by the fact that repeated questioning, expenditure of effort over time, and public displays of confidence (as might happen at a trial) all tend to inflate eyewitness confidence even when accuracy is held constant (Shaw & McClure, 1996; Shaw & Zerr, 2003; Shaw, Zerr, & Woythaler, 2001). Clearly, the most pristine measure of witness confidence is one collected from the witness at the time of identification and before the contaminating influence of these later events.

An intriguing phenomenon that appears to be related to the post-identification feedback effect is *visual hindsight bias*. Harley, Carlsen, and Loftus (2004) presented participants with photographs of familiar faces that were severely degraded (blurred) but gradually resolved to full clarity. After the identity of the face became apparent, participants predicted the level of blur that would permit a naïve observer to identify the face. Participants who had already learned the identity of the face consistently predicted that a naïve participant would be able to identify the face at levels of blur that were in fact too severe for identification. Thus, once the "correct" answer is known, people think that objectively poor viewing conditions are nevertheless sufficient for accurate identification. This "saw it all along" effect could be an important component of the propensity for eyewitnesses to have retrospective overconfidence in their identifications.

Response Latency

Another interesting postdictor of eyewitness accuracy is the *response latency* of the eyewitness in making a lineup identification. We use "response latency" rather than "decision time," because the former term incorporates both decisional and motor components (Weber, Brewer, Wells, Semmler, & Keast, 2004). The effect was first documented by Sporer (1992); considerable data have accumulated showing that witnesses who make accurate identifications from lineups do so faster than do those who make inaccurate identifications (Dunning & Perretta, 2002; Dunning & Stern, 1994; Smith, Lindsay, & Pryke, 2000; Smith, Lindsay, Pryke, & Dysart, 2001; Sporer, 1993, 1994; Weber et al., 2004). Sporer (1992) suggested that this occurs because comparisons made to the target involve a large number of common features between memory and the stimulus face, thereby permitting a very fast decision in recognizing the target. Comparisons to an innocent lineup member, on the other hand, involve fewer common features between memory and the stimulus, thereby resulting in a slower decision. The potential practical value of the negative relation between response latency and identification accuracy is considerable because, unlike confidence, response latency is a performance variable rather than a self-report. And, unlike confidence, response latency can be measured without the eyewitness's awareness. Furthermore, response latency and confidence are not fully redundant postdictors of accuracy (Smith et al., 2001; Weber et al., 2004).

For response latency to be useful at the level of evaluating an individual eyewitness, however, some criteria have to be set for "fast" and "slow." How are police, prosecutors, judges, and juries to know whether a given result (e.g., response latency of 20 seconds) was fast or slow and thus should be considered accurate or inaccurate? Dunning and Perretta (2002) approached this problem by repeatedly selecting different response latencies, examining the percentages correct above and below each response latency, and calculating the obtained chi-square values for each response latency. The response latency that produced the greatest value was then considered to be the best rule for deciding on the decision criterion. Using this approach, Dunning and Perretta found that a response latency of 10 to 12 seconds worked best across four different data sets. Furthermore, the 10–12-second response latency was highly discriminating—those who responded before the 10–12-second latency had a probability of accuracy of nearly 90%, while those who took longer than 10–12 seconds had a probability of accuracy of approximately 50%. Dunning and Perretta called this the "10–12 second rule." The consistency of the 10–12-second response latency data sets fits nicely with Dunning and Stern's (1994) notion of automatic versus deliberative processing in eyewitness identification. They argued that automatic decision processes (which are fast) are likely to be characteristic of accurate eyewitnesses, while deliberative processes (which are slower) ought to be more characteristic of inaccurate eyewitnesses. Furthermore, because automatic processes tend to be uninfluenced by decision context, the speed of accurate identifications ought to be relatively stable across situations—hence, the 10–12-second rule was proposed to be stable across various circumstances and conditions.

More recent research, however, has shown that the 10–12-second rule is not stable across variations in witnessing and lineup conditions. Weber et al. (2004) found that the maximally discriminating time ranged from 5 seconds to 29 seconds across variations in conditions. Furthermore, eyewitnesses who responded faster than the optimal time boundaries did not show particularly high probabilities of being accurate; they were often in the 50 to 60% range rather than the 90% range found by Dunning and Perretta (2002). Although the 10–12-second rule does not appear to be stable, the fact that accurate identifications are made faster than inaccurate identifications is itself a very reliable finding.

Self-Reported Decision Processes

Another potential postdictor of eyewitness-identification accuracy is eyewitnesses' reports of the processes they use to make their identification decisions. Wells (1984) argued that mistaken identifications tend to arise from making relative judgments in which the eyewitness compares one lineup member to another to decide who looks most like the target; Wells argued that an absolute judgment (comparing the lineup member to memory) would be more likely to be accurate. Consistent with this assumption, Stern and Dunning (1994) found that eyewitnesses who agreed with the statement "I compared the photos [in the lineup] to each other to narrow the choices" were more likely to have made a mistaken identification than were those who endorsed the statement "I just recognized him, I cannot explain why" or those who said the photo "popped out." Similar results have been reported by Smith et al. (2000), Smith et al. (2001), Dunning and Stern (1994), and Lindsay and Bellinger (1999).

One of the problems with self-reported decision processes is that, like eyewitness confidence, they are subject to distortion. For instance, confirmatory post-identification feedback leads eyewitnesses to be more likely to recall that the lineup photo "popped out" and less likely to report having made a relative judgment (Wells & Bradfield, 1998). Furthermore, if eyewitnesses thought these kinds of self-reports would be used to assess the likely accuracy of their identifications, they might shape their answers accordingly.

Overall, it appears that postdiction has not been highly successful for eyewitness identification. Indicators of confidence measured at the time of the identification may have some diagnostic value with regard to accuracy, but feedback, prosecutorial skimming, and plea bargaining can operate to obliterate the diagnostic value of confidence. This underscores the primary message of the system-variable approach—namely, that it would be better to use procedures that help prevent mistaken identifications from occurring in the first place than to try to detect errors after the fact.

PROGRESS AND PROSPECTS

Eyewitness science has made considerable progress in recent years in getting a number of jurisdictions in the United States to improve their identification procedures and undertake training in the cognitive interview. The state of New Jersey, for instance, has adopted an entire package of reforms for how it conducts lineups. These reforms are based explicitly on the eyewitness literature and include the adoption of recommendations for selecting lineup fillers, instructing eyewitnesses before the lineup, using double-blind lineup administrators, using the sequential procedure, and obtaining a confidence statement from the eyewitness before external factors can influence the person's confidence. Other jurisdictions—including the states of Wisconsin and North Carolina and the cities of Boston and Minneapolis—have also adopted these reforms. In each of these jurisdictions, eyewitness scientists played a central role in explaining the literature and helping translate the findings into practical reforms of eyewitness-identification procedures.

In many jurisdictions, eyewitness researchers have become involved in training police investigators in eyewitness-identification procedures or training the trainers. Increasingly, eyewitness researchers are targeting some of their writings toward law enforcement journals to more directly make the research findings accessible to law enforcement (e.g., Turtle, Lindsay, & Wells, 2003). Jury simulations have shown that mock jurors respond more favorably to eyewitness-identification testimony when it was obtained using these packages of reformed procedures than when procedures deviate from these reforms (Lampinen, Judges, Odegard, & Hamilton, 2005). This is an unusual impact for a laboratory-based psychological science. In the years ahead, it is expected that these reforms will become even more widespread and the role of scientific psychology more deeply ingrained in the legal system.

Despite this progress, we believe that research has only scratched the surface of ways to help the legal system improve the accuracy of eyewitness accounts. Thus far, almost all improvements to lineup procedure have been designed to reduce the chances that an innocent suspect will be identified without reducing identifications of the target. It has been more difficult to discover ways to increase the chances that the eyewitness will identify the target in target-present lineups. Both research experiments and archival analyses of actual lineups suggest that eyewitnesses fail to identify the target about 50% of the time. This does not necessarily mean that the target walks away; in some cases, other evidence is sufficient to charge or convict the person. Nevertheless, there is room to improve these hit rates. It seems likely that some failures to identify the target from target-present lineups are due at least in part to changes in the target's appearance. Specifically, the appearance of the target when the eyewitness viewed the crime represents a moment in time. The photo seen in a photographic lineup may be older or more recent. Attempts to use pre-lineup instructions that warn the eyewitness that the target's appearance might have changed have not proved successful in increasing accuracy; in fact, they seem to increase errors (Charman & Wells, in press).

It could be argued that research has been profoundly conservative in its approach to the eyewitness-identification problem. Specifically, researchers have tended to operate within the confines of the traditional lineup, in which a suspect is placed among fillers and the eyewitness makes a verbal identification. But what if the lineup had never existed and the legal system turned to psychology to determine how information could be extracted from eyewitnesses' memories? Specific methods for obtaining detailed reports from witnesses—such as the cognitive interview—do not appear to aid identification, but the quality of witness descriptions could be improved though innovative questioning procedures. This is an area in which research is sparse, despite the potential to study the effectiveness of various types of retrieval cues in eliciting descriptions (Sporer, in press). The focus on target identification has also resulted in research that has selectively focused on the impact of a specific system or estimator variable on lineup performance, instead of exploring relevant interactions. For example, is the weapon-focus effect more pronounced when a witness has a shorter exposure to the target, when the retention interval is longer, and when the witness is making a cross-race identification? Operating from scratch, it seems likely that modern psychology would have developed radically different ideas. For instance, brain-activity measures, eye movements, rapid displays of faces, reaction times, and other methods for studying memory might have been developed instead of the traditional lineup. Once we step outside the confines of the traditional lineup, it is possible to imagine a future science of eyewitness evidence that is radically different from the methods used today.

REFERENCES

Akehurst, L., Milne, R., & Köhnken, G. (2003). The effects of children's age and delay on recall in a cognitive or structured interview. *Psychology, Crime and Law*, *9*, 97–107.

Allport, G., & Postman, L. (1947). *The psychology of rumor*. New York: Holt.

Ayers, M.S., & Reder, L.M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based model. *Psychonomic Bulletin and Review*, *5*, 1–21.

Bekerian, D.A., & Dennett, J.L. (1997). Imagery effects in spoken and written recall. In D. Payne & F. Conrad (Eds.), *Intersections in basic and applied memory research* (pp. 175–191). Mahwah, NJ: Erlbaum.

Behrman, B.W., & Davey, S.L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*, 475–491.

Berman, G.L., & Cutler, B.L. (1996). Effects of inconsistencies in eyewitness testimony on mock-juror decision making. *Journal of Applied Psychology*, *81*, 170–177.

Berresheim, A., & Weber, A. (2003). Die Strukturierte Zeugenvernehmung und Ihre Wirksamkeit [Structured witness interviewing and its effectiveness]. *Kriminalistik*, *57*, 757–771.

Binet, A. (1900). *La suggestibilité* [On suggestibility]. Paris: Schleicher.

Boon, J., & Noon, E. (1994). Changing perspectives in cognitive interviewing. *Psychology, Crime and Law*, *1*, 59–69.

Borchard, E. (1932). *Convicting the innocent: Errors of criminal justice*. New Haven, CT: Yale University Press.

Bornstein, B., & Zickafoose, D.J. (1999). "I know I know it, I know I saw it": The stability of the confidence-accuracy relationship across domains. *Journal of Experimental Psychology: Applied*, *5*, 76–88.

Bower, G. (1967). A multicomponent theory of memory trace. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 229–325). New York: Academic Press.

Bradfield, A.L., & Wells, G.L. (2000). The perceived validity of eyewitness identification testimony: A test of the five Biggers criteria. *Law and Human Behavior*, *24*, 581–594.

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.

Brainerd, C.J., & Reyna, V.F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129.

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*, 46–58.

Brewer, N., Weber, N., & Semmler, C. (2005). Eyewitness identification. In N. Brewer & K.D. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 177–221). New York: Guilford.

Brewer, N., & Wells, G.L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, filler similarity and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*, 11–30.

Brigham, J.C. (1990). Target person distinctiveness and attractiveness as moderator variables in the confidence-accuracy relationship in eyewitness identifications. *Basic & Applied Social Psychology*, *11*, 101–115.

Brigham, J.C., & Bothwell, R.K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identifications. *Law and Human Behavior*, *7*, 19–30.

Brigham, J.C., Maas, A., Snyder, L.D., & Spaulding, K. (1982). Accuracy of eyewitness identifications in a field setting. *Journal of Personality and Social Psychology*, *42*, 673–681.

Brigham, J.C., Meissner, C.A., & Wasserman, A.W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology*, *13*, S73–S92.

Bruce, V., Ness, H., Hancock, P.J.B., Newman, C., & Rarity, J. (2002). Four heads are better than one: Combining face composites yields improvements in face likeness. *Journal of Applied Psychology*, *87*, 894–902.

Bruck, M., & Ceci, S.J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419–439.

Buckhout, R. (1974). Eyewitness testimony. *Scientific American*, *231*, 23–31.

Buckhout, R. (1980). Nearly 2,000 witnesses can be wrong. *Bulletin of the Psychonomic Society*, *4*, 191–192.

Bull, R., & Clifford, B.R. (1976). Identification: The Devlin Report. *New Scientist*, *70*, 307–308.

Burtt, H. (1931). *Legal psychology*. Englewood Cliffs, NJ: Prentice Hall.

Charman, S.D., & Wells, G.L. (2006). Applied lineup theory. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (pp. 219–254). Mahwah, NJ: Erlbaum.

Charman, S.D., & Wells, G.L. (in press). Eyewitness lineups: Is the appearance-change instruction a good idea? *Law and Human Behavior*.

Clarke, C., & Milne, R. (2001). National Evaluation of the PEACE investigative interviewing course. Police Research Award Scheme (Report No. PRAS/149). London: Home Office.

Clark, S.E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology*, *17*, 629–654.
Clark, S.E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, *29*, 395–424.
Clark, S.E., & Davey, S.L. (2005). The target-to-fillers shift in simultaneous and sequential lineups. *Law and Human Behavior*, *29*, 151–172.
Clark, S.E., & Tunnicliff, J.L. (2001). Selecting lineup fillers in eyewitness identification: Experimental control and real-world simulation. *Law and Human Behavior*, *25*, 199–216.
Clifford, BR., & Bull, R. (1978). *The psychology of person identification*. London: Routledge & Kegan Paul.
Clifford, B.R., & Hollin, C.R. (1981). Effects of the type of incident and the number of perpetrators on eyewitness memory. *Journal of Applied Psychology*, *66*, 364–370.
Clifford, B.R., & Scott, J. (1978). Individual and situational factors in eyewitness testimony. *Journal of Applied Psychology*, *63*, 352–359.
Cole, G.F. (1986). *The American system of criminal justice* (4th ed.). Monterey, CA: Brooks/Cole.
Connors, E., Lundregan, T., Miller, N., & McEwan, T. (1996). *Convicted by juries, exonerated by science: Case studies in the use of DNA evidence to establish innocence after trial*. Alexandria, VA: National Institute of Justice.
Cote, P. (1998). *FACES:* The Ultimate Composite Picture (Version 3.0) [Computer software]. Saint Hubert, Quebec, Canada: Interquest.
Cutler, B.L., & Penrod, S.D. (1989). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology*, *74*, 650–653.
Cutler, B.L., & Penrod, S.D. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. New York: Cambridge University Press.
Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987a). Improving the reliability of eyewitness identifications: Putting context into context. *Journal of Applied Psychology*, *72*, 629–637.
Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987b). The reliability of eyewitness identifications: The role of system and estimator variables. *Law and Human Behavior*, *11*, 223–258.
Cutler, B.L., Penrod, S.D., & Stuve, T.E. (1988). Jury decision making in eyewitness identification cases. *Law and Human Behavior*, *12*, 41–56.
Davies, G.M., Ellis, H.D., & Shepherd, J.W. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, *63*, 180–187.
Davis, M., McMahon, M., & Greenwood, K. (2004). The efficacy of the mnemonic components of the cognitive interview: Towards a shortened variant for time-critical investigations. *Applied Cognitive Psychology*, *19*, 75–93.
Deffenbacher, K. (1980). Eyewitness accuracy and confidence: Can we infer anything about their relationship? *Law and Human Behavior*, *4*, 243–260.
Deffenbacher, K.A., Bornstein, B.H., Penrod, S.D., & McGorty, E.K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, *28*, 687–706.
Devlin, L. (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases*. London: Her Majesty's Stationery Office.
Dixon, S., & Memon, A. (2005). The effect of post-identification feedback on the recall of crime and perpetrator details. *Applied Cognitive Psychology*, *19*, 935–951.
Doob, A.N., & Kirshenbaum, H. (1973). Bias in police lineups—Partial remembering. *Journal of Police Science and Administration*, *1*, 287–293.
Douglass, A.B., & Steblay, N. (in press). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology*.
Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy: A 10–12 second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology*, *87*, 951–962.
Dunning, D., & Stern, L.B. (1994). Distinguishing accurate from inaccurate identifications via inquiries about decision processes. *Journal of Personality and Social Psychology*, *67*, 818–835.
Dysart, J.E., Lindsay, R.C.L., MacDonald, T.K., & Wicke, C. (2002). The intoxicated witness: Effects of alcohol on identification accuracy from showups. *Journal of Applied Psychology*, *87*, 170–175.
Ebbeson, E.B., & Flowe, H.D. (n.d.). *Simultaneous v. sequential lineups: What do we really know*? Retrieved October 12, 2006, from http://www-psy.ucsd.edu/~eebbesen/SimSeq.htm#_edn1
Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition*, *30*, 770–782.
Ellis, H.D., Davies, G.M., & Shepherd, J.W. (1977). Experimental studies of face identification. *Journal of Criminal Defense*, *3*, 219–234.
Ellis, H.D., Davies, G.M., & Shepherd, J.W. (1978). A critical examination of the Photo-fit system for recalling faces. *Ergonomics*, *21*, 297–301.
Ellis, H.D., Shepherd, J.W., Flin, R.H., Shepherd, J., & Davies, G.M. (1989). Identification from a computer-driven retrieval system compared with traditional mugshot album search: A new tool for police investigations. *Ergonomics*, *32*, 167–177.
Fisher, R.P. (2005). *The cognitive interview: Question & answers*. Ashburn, Virginia: The National Transportation Safety Board Academy. Retrieved June 13, 2006, from http://www.ntsb.gov/Academy/courseinfo/Q&A_Fisher.pdf
Fisher, R.P., Brennan, K.H., & McCauley, M.R. (2002). The cognitive interview method to enhance eyewitness recall. In M. Eisen, G. Goodman, & J. Quas (Eds.), *Memory and suggestibility in the forensic interview* (pp. 265–286). Mahwah, NJ: Erlbaum.
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 interviewing techniques. *Journal of Police Science and Administration*, *15*, 177–185.
Fisher, R.P., Quigley, K.L., Brock, P., Chin, E., & Cutler, B.L. (1990, March). *The effectiveness of the cognitive interview in description and identification tasks*. Paper presented at the American Psychology-Law Society, Williamsburg, VA.
Fisher, R.P., & Schreiber, N. (in press). Interviewing protocols to improve eyewitness memory. In M. Toglia, R. Lindsay, D. Ross, & J. Reed (Eds.), *The handbook of eyewitness psychology: Vol. 1. Memory for events*. Mahwah, NJ: Erlbaum.
Fleet, M.L., Brigham, J.C., & Bothwell, R.K. (1987). The confidence-accuracy relationship: The effects of confidence assessment and choosing. *Journal of Applied Social Psychology*, *17*, 171–187.
Fox, S.G., & Walters, H.A. (1986). The impact of general versus specific expert testimony and eyewitness confidence upon mock juror judgment. *Law and Human Behavior*, *10*, 215–228.

Frank, J., & Frank, B. (1957). *Not guilty*. London: Gallanez.

Frowd, C.D., Carson, D., Ness, H., McQuiston, D., Richardson, J., Baldwin, H., & Hancock, P.J.B. (2005). Contemporary composite techniques: The impact of a forensically relevant target delay. *Legal and Criminological Psychology*, *10*, 63–81.

Gabbert, F., Memon, A., & Wright, D.B. (in press). Memory conformity: Disentangling the steps towards influence during a discussion. *Psychonomic Bulletin and Review*.

Garrioch, L., & Brimacombe, C.A.E. (2001). Lineup administrators' expectations: Their impact on eyewitness confidence. *Law and Human Behavior*, *25*, 299–314.

Geiselman, R.E., & Callot, R. (1990). Reverse versus forward recall of script-based texts. *Applied Cognitive Psychology*, *4*, 141–144.

Geiselman, R.E., Fisher, R.P., Firstenberg, I., Hutton, L.A., Sullivan, S.J., Avetissian, I.V., & Prosk, A.L. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science and Administration*, *12*, 74–79.

Geiselman, R.E., Fisher, R.P., MacKinnon, D.P., & Holland, H.L. (1985). Eyewitness memory enhancement in the police interview. *Journal of Applied Psychology*, *27*, 358–418.

George, R.C. (1991). *A field evaluation of the cognitive interview*. Unpublished master's thesis, Polytechnic of East London.

Granhag, PA., & Stromwall, L. (2004). *The detection of deception in forensic contexts*. Cambridge, England: Cambridge University Press.

Gwyer, P., & Clifford, B.R. (1997). The effects of the cognitive interview on recall, identification, confidence and the confidence/accuracy relationship. *Applied Cognitive Psychology*, *11*, 121–145.

Hafstad, G.S., Memon, A., & Logie, R. (2004). Post-identification feedback, confidence, and recollections of witnessing conditions in child witnesses. *Applied Cognitive Psychology*, *18*, 901–912.

Harley, E.M., Carlsen, K.A., & Loftus, G.R. (2004). The "saw-it-all-along" effect: Demonstrations of visual hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 960–968.

Harris, M.J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, *97*, 363–386.

Hastorf, A.H., & Cantrill, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology*, *49*, 129–134.

Hasel, L., & Wells, G.L. (in press). Catching the bad guy: Morphing composite faces helps. *Law and Human Behavior*.

Haw, R.M., & Fisher, R.P. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology*, *89*, 1106–1112.

Holliday, R.E. (2003a). Reducing misinformation effects in children with cognitive interviews: Dissociating recollection and familiarity. *Child Development*, *74*, 728–751.

Holliday, R.E. (2003b). The effect of a prior cognitive interview on children's acceptance of misinformation. *Applied Cognitive Psychology*, *17*, 443–457.

Hosch, H.M., & Cooper, D.S. (1982). Victimization as a determinant of eyewitness accuracy. *Journal of Applied Psychology*, *67*, 649–652.

Hosch, H.M., Leippe, M.R., Marchioni, P.M., & Cooper, D.S. (1984). Victimization, self-monitoring, and eyewitness identification. *Journal of Applied Psychology*, *69*, 280–288.

Huff, R., Rattner, A., & Sagarin, E. (1986). Guilty until proven innocent. *Crime and Delinquency*, *32*, 518–544.

Illinois Pilot Program on Double-Blind Sequential Procedures. (2006). Retrieved March 21, 2006, from http://www.chicagopolice.org/IL%20Pilot%20on%20Eyewitness%20ID.pdf

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can and cannot be inferred from a low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *5*, 1304–1316.

Kassin, S.M., Tubb, V.A., Hosch, H.M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research. *American Psychologist*, *56*, 405–416.

Kassin, S.M., & Gudjonsson, G.H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33–67.

Kebbell, M., & Milne, R. (1998). Police officers' perception of eyewitness factors in forensic investigations. *Journal of Social Psychology*, *138*, 323–330.

Kebbell, M.R., Milne, R., & Wagstaff, G. (1999). The cognitive interview: A survey of its forensic effectiveness. *Psychology, Crime and Law*, *5*, 101–115.

Kebbell, M.R., Milne, R., & Wagstaff, G. (2001). The cognitive interview in forensic investigations: A review. In G.B. Traverson & L. Bagnoli (Eds.), *Psychology and law in a changing world: New trends in theory, research and practice* (pp. 185–198). London: Routledge.

Klobuchar, A., Steblay, N., & Caligiuri, H. (in press). Improving eyewitness identifications: Hennepin County's blind sequential lineup project. *Cardozo Public Law, Policy, and Ethics Journal*.

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). A meta-analysis on the effects of the cognitive interview. *Psychology, Crime and Law*, *5*, 3–27.

Köhnken, G., Schimmossek, E., Aschermann, E., & Höfer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology*, *80*, 671–684.

Köhnken, G., Thurer, C., & Zorberbier, D. (1994). The cognitive interview: Are interviewers' memories enhanced too? *Applied Cognitive Psychology*, *8*, 13–24.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Towards a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481–537.

Kovera, M.B., Penrod, S.D., Pappas, C., & Thill, D.L. (1997). Identification of computer-generated facial composites. *Journal of Applied Psychology*, *82*, 235–246.

Lampinen, J.M., Judges, D.P., Odegard, T.N., & Hamilton, S. (2005). The reactions of mock jurors to the Department of Justice Guidelines for the Collection and Preservation of Eyewitness Evidence. *Basic and Applied Social Psychology*, *27*, 155–162.

Leippe, M.R. (1980). Effect of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior*, *4*, 261–274.

Levi, A.M. (2002). Lineup size, the modified sequential lineup, and the sequential lineup. *Cognitive Technology*, *7*, 39–46.

Lindsay, D.S., & Johnson, M.K. (1989). The eyewitness suggestibility effect and memory for source. *Memory and Cognition*, *17*, 349–358.

Lindsay, R.C.L. (1986). Confidence and accuracy in eyewitness identification from lineups. *Law and Human Behavior*, *10*, 229–239.

Lindsay, R.C.L., & Bellinger, K. (1999). Alternatives to the sequential lineup: The importance of controlling the pictures. *Journal of Applied Psychology*, *84*, 315–321.

Lindsay, R.C.L., Lim, R., Mirando, L., & Cully, D. (1986). Mock juror evaluations of eyewitness testimony: A test of metamemory hypotheses. *Journal of Applied Social Psychology*, *16*, 447–459.

Lindsay, R.C.L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup filler selection strategy. *Law and Human Behavior*, *18*, 527–541.

Lindsay, R.C.L., Nosworthy, G.J., Martin, R., & Martynuck, C. (1994). Finding suspects in mugshots. *Journal of Applied Psychology*, *79*, 121–130.

Lindsay, R.C.L., & Wells, G.L. (1980). What price justice? Exploring the relationship between lineup fairness and identification accuracy. *Law and Human Behavior*, *4*, 303–314.

Lindsay, R.C.L., & Wells, G.L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentations. *Journal of Applied Psychology*, *70*, 556–564.

Lindsay, R.C.L., Wells, G.L., & O'Connor, F. (1989). Mock juror belief of accurate and inaccurate eyewitnesses: A replication. *Law and Human Behavior*, *13*, 333–340.

Lindsay, R.C.L., Wells, G.L., & Rumpel, C. (1981). Can people detect eyewitness identification accuracy within and between situations? *Journal of Applied Psychology*, *66*, 79–89.

Loftus, E.F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E.F. (2005). A 30-year investigation of the malleability of memory. *Learning and Memory*, *12*, 361–366.

Loftus, E.F., & Palmer, J.C. (1974). Reconstruction of an automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.

Luus, C.A.E., & Wells, G.L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior*, *15*, 43–57.

Luus, C.A.E., & Wells, G.L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, *79*, 714–724.

Maas, A., Brigham, J., & West, S.G. (1985). Testifying on eyewitness reliability: Expert advice is not always persuasive. *Journal of Applied Social Psychology*, *15*, 207–229.

MacLeod, M. (2002). Retrieval-induced forgetting in eyewitness memory: Forgetting as a consequence of remembering. *Applied Cognitive Psychology*, *16*, 135–149.

MacLin, O.H., Zimmerman, L.A., & Malpass, R.S. (2005). PC_Eyewitness and the sequential superiority effect: Computer-based lineup administration. *Law and Human Behavior*, *29*, 303–321.

MacLin, O., Wells, G.L., & Phelan, C. (2004). [Target removal-without-replacement in lineups]. Unpublished raw data.

Malpass, R.S., & Devine, P.G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, *66*, 482–489.

Malpass, R.S., & Lindsay, R.C.L. (1999). Measuring lineup fairness. *Applied Cognitive Psychology*, *13*(Suppl. 1), S1–S8.

Manson v. Braithwaite, 432 U.S. 98 (1977).

McCloskey, M., & Egeth, H. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist*, *38*, 550–563.

McQuiston, D.E., & Malpass, R.S. (2002). Validity of the mock witness paradigm: Testing the assumptions. *Law and Human Behavior*, *26*, 439–453.

McKelvie, S.J. (1988). The role of spectacles in facial memory: A replication and extension. *Perceptual and Motor Skills*, *66*, 651–658.

Meissner, C.A., & Brigham, J.C. (2001a). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*, 3–35.

Meissner, C.A, & Brigham, J.C. (2001b). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology*, *15*, 603–616.

Meissner, C.A., Sporer, S.L., & Schooler, J.W. (in press). Person descriptions as eyewitness evidence. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (pp. 1–34). Mahwah, NJ: Erlbaum.

Meissner, C.A., Tredoux, C.J., Parker, J.F., & MacLin, O.H. (2005). Eyewitness decisions in simultaneous and sequential lineups. *Memory and Cognition*, *33*, 783–792.

Mello, E.W., & Fisher, R.P. (1996). Enhancing older adult eyewitness memory with the cognitive interview. *Applied Cognitive Psychology*, *10*, 403–418.

Memon, A. (2006). The cognitive interview. In O. Hargie (Ed.), *A handbook of communication skills* (3rd ed.). London and New York: Routledge.

Memon, A., & Bull, R. (Eds.). (1999). *Handbook of the psychology of interviewing*. Chichester, England: Wiley.

Memon, A., Cronin, Ó., Eaves, R., & Bull, R. (1993). The cognitive interview and child interviewees. In G.M. Stephenson & N.K. Clark (Eds.), *Children, evidence and procedure: Issues in criminological & legal psychology*. Division of Criminological and Legal Psychology (Occasional Paper No. 20). Leicester, England: British Psychological Society.

Memon, A., & Higham, P. (1999). A review of the cognitive interview. *Psychology, Crime and Law*, *5*, 177–196.

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 and Cognition*, *30*, 1219–1227.

Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology*, *94*, 339–354.

Memon, A., Milne, R., Holley, A., Bull, R., & Köhnken, G. (1994). Towards understanding the effects of interviewer training in evaluating the cognitive interview. *Applied Cognitive Psychology*, *8*, 641–659.

Memon, A., & Stevenage, S.V. (1996). Interviewing witnesses: What works and what doesn't? *Psycholoquy* 7(6), Witness Memory (1). Retrieved October 20, 2006, from http://psycprints.ecs.soton.ac.uk/archive/00000498/

Memon, A., Wark, L., Bull, R., & Köhnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology*, *88*, 179–198.

Memon, A., Wark, L., Holley, A., Bull, R., & Köhnken, G. (1997). Eyewitness performance in cognitive and structured interviews. *Memory*, *5*, 639–655.

Milne, R., & Bull, R. (2006). Interviewing victims of crime, including children and people with intellectual difficulties. In M.R. Kebbell & G.M. Davies (Eds.), *Practical psychology for forensic investigations*. Chichester, England: Wiley.

Milne, R., Bull, R., Memon, A., & Köhnken, G. (1995). The cognitive interview and suggestibility. In G.M. Stephenson & N.K. Clark (Eds.), *Criminal behaviour: Perceptions, attributions and rationality*. Division of Criminological and Legal Psychology (Occasional Paper No. 22). Leicester, England: British Psychological Society.

Milne, R., Clare, I.C.H., & Bull, R. (1999). Using the cognitive interview with adults with mild learning disabilities. *Psychology, Crime and Law*, *5*, 81–100.

Mitchell, K.J., & Johnson, M. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179–195). New York: Oxford University Press.

Mitchell, K.J., Livosky, M., & Mather, M. (1998). The weapon focus effect revisited: The role of novelty. *Legal and Criminological Psychology*, *3*, 287–303.

Morgan, C.A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baronoski, M., & Southwick, S.M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry*, *27*, 265–279.

Münsterberg, H. (1908). *On the witness stand*. New York: Doubleday.

Newsome v. McCabe et al., 256 F.3d 747 (2002).

Neuschatz, J.S., Preston, E.L., Burkett, A.D., Toglia, M.R., Lampinen, J.M., Neuschatz, J.S., Fairless, A.H., Lawson, D.S., Powers, R.A., & Goodsell, C.A. (2005). The effects of post-identification feedback and age on retrospective eyewitness memory. *Applied Cognitive Psychology*, *19*, 435–453.

Patterson, K.E., & Baddeley, A.D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 406–417.

Pecher, D., Zeelenberg, R., & Barsalou, L.W. (2003). Verifying conceptual properties in different modalities produces switching costs. *Psychological Science*, *14*, 119–124.

Penrod, S. (2003, Spring). Eyewitness identification evidence: How well are witnesses and police performing? *Criminal Justice*, *54*, 36–47.

Perfect, T.J., & Hollins, T.S. (1996). Predictive feeling of knowing judgments and postdictive confidence judgments in eyewitness memory and general knowledge. *Applied Cognitive Psychology*, *10*, 371–382.

Pezdek, K., Blandon-Gitlin, I., & Moore, C. (2003). Children's face recognition memory: More evidence for the cross-race effect. *Journal of Applied Psychology*, *88*, 760–763.

Phillips, M.R., McAuliff, B.D., Kovera, M.B., & Cutler, B.L. (1999). Double-blind photoarray administration as a safeguard against investigator bias. *Journal of Applied Psychology*, *84*, 940–951.

Pickel, K.L. (1998). Unusualness and threat as possible causes of "weapon focus."*Memory*, *6*, 277–295.

Pickel, K.L. (1999). The influence of context on the "weapon focus" effect. *Law and Human Behavior*, *23*, 299–311.

Pike, G.E., Kemp, R.I., Towell, N.A., & Phillips, K.C. (1997). Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition*, *4*, 409–437.

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

Poole, D., & Lindsay, D.S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied*, *7*, 27–50.

Read, J.D., Vokey, J.R., & Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 870–882.

Read, J.D., Yuille, J.C., & Tollestrup, P. (1992). Recollections of a robbery: Effects of arousal and alcohol upon recall and person identification. *Law and Human Behavior*, *16*, 425–446.

Roberts, W.T., & Higham, P.A. (2002). Selecting accurate statements from the cognitive interview using confidence ratings. *Journal of Experimental Psychology: Applied*, *8*, 33–43.

Roberts, K., Lamb, M., & Sternberg, K. (2004). The effects of rapport building style on children's reports of a staged event. *Applied Cognitive Psychology*, *18*, 189–202.

Robinson, M.D., & Johnson, J.T. (1998). How not to enhance the confidence-accuracy relation: The detrimental effects of attention to the identification process. *Law and Human Behavior*, *22*, 409–428.

Savage, S., & Milne, R. (in press). Miscarriages of justice—the role of the investigative process. In T. Newburn, T. Williamson, & A. Wright (Eds.), *Handbook of criminal investigation*. Cullompton, England: Willan.

Saywitz, K.J., Geiselman, R.E., & Bornstein, G.K. (1992). Effects of cognitive interviewing and practice on children's recall performance. *Journal of Applied Psychology*, *77*, 744–756.

Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence*. New York: Random House.

Schooler, J.W., & Engstler-Schooler, T.Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*, 36–71.

Searcy, J.H., Bartlett, J.C., & Memon, A. (1999). Age differences in accuracy and choosing rates on face recognition and eyewitness identification tasks. *Memory and Cognition*, *27*, 538–552.

Searcy, J.H., Bartlett, J.C., Swanson, K., & Memon, A. (2001). Aging and lineup performance at long retention intervals: Effects of meta-memory and context reinstatement. *Journal of Applied Psychology*, *86*, 207–214.

Semmler, C., Brewer, N., & Wells, G.L. (2004). Effects of postidentification feedback on eyewitness identification and nonidentification. *Journal of Applied Psychology*, *89*, 334–346.

Shapiro, P., & Penrod, S. (1986). A meta-analysis of facial identification studies. *Psychological Bulletin*, *100*, 139–156.

Shaw, J.S., III, & McClure, K.A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior*, *20*, 629–654.

Shaw, J.S., III, & Skolnick, P. (1999). Weapon focus and gender differences in eyewitness accuracy: Arousal versus salience. *Journal of Applied Social Psychology*, *29*, 2328–2341.

Shaw, J.S., III, & Zerr, T.K. (2003). Extra effort during memory retrieval may be associated with increases in eyewitness confidence. *Law and Human Behavior*, *27*, 315–329.

Shaw, J.S., III, Zerr, T.K., & Woythaler, K.A. (2001). Public eyewitness confidence ratings can differ from those held privately. *Law and Human Behavior*, *25*, 141–154.

Slater, A. (1994). *Identification parades: A scientific evaluation*. London: Police Research Group, Home Office.

Smith, S.M. (1994). Theoretical principles of context dependent memory. In P. Morris & M. Grunerberg (Eds.), *Aspects of memory: Theoretical aspects* (2nd ed., pp. 169–195). New York: Routledge Press.

Smith, S.M., Lindsay, R.C.L., & Pryke, S. (2000). Postdictors of eyewitness errors: Can false identification be diagnosed? *Journal of Applied Psychology*, *85*, 542–550.

Smith, S.M., Lindsay, R.C.L., Pryke, S., & Dysart, J.E. (2001). Postdictors of eyewitness errors: Can false identifications be diagnosed in the cross-race situation? *Psychology, Public Policy, and Law*, *7*, 153–169.

Smith, S.M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin and Review*, *8*, 203–220.

Snee, T., & Lush, D. (1941). Interaction of the narrative and interrogatory methods of obtaining testimony. *Journal of Psychology*, *11*, 229–230.

Sporer, S.L. (1982). A brief history of the psychology of testimony. *Current Psychological Reviews*, *2*, 323–339.

Sporer, S.L. (1992). Post-dicting eyewitness identification accuracy: Confidence, decision-times and person descriptions among choosers and nonchoosers. *European Journal of Social Psychology*, *20*, 157–180.

Sporer, S.L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology*, *78*, 22–33.

Sporer, S.L. (1994). Decision-times and eyewitness identification accuracy in simultaneous and sequential lineups. In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 300–327). New York: Cambridge University Press.

Sporer, S.L. (1996). Psychological aspects of person descriptions. In S. Sporer, R. Malpass, & G. Köhnken (Eds.), *Psychological issues in eyewitness identification* (pp. 53–86). Mahwah, NJ: Erlbaum.

Sporer, S.L. (in press). Person descriptions as retrieval cues: Do they really help? *Psychology, Crime and Law.*

Sporer, S., Penrod, S., Read, D., & Cutler, B.L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*, 315–327.

State of Maryland v. Kirk N. Bloodsworth, 84-CR-3138 (Baltimore County 1984).

Steblay, N.M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, *16*, 413–424.

Steblay, N.M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, *21*, 283–298.

Steblay, N.M., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *25*, 459–474.

Steele, C.M., & Josephs, R.A. (1990). Alcohol myopia: Its prized and dangerous effects. *American Psychologist*, *45*, 921–933.

Stein, L., & Memon, A. (in press). Testing the efficacy of the cognitive interview in a developing country. *Applied Cognitive Psychology.*

Stern, L.B., & Dunning, D. (1994). Distinguishing accurate from inaccurate eyewitness identifications: A reality monitoring approach. In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 273–299). New York: Cambridge University Press.

Stern, L.W. (1904). Die Aussage als geistige Leistung und als Verhörsprodukt [The statement as a mental achievement and product of interrogation]. *Beiträge zur Psychologie der Aussage*, *3*, 269–415.

Stern, L.W. (1939). The psychology of testimony. *Journal of Abnormal and Social Psychology*, *34*, 3–20.

Stevenage, S., & Memon, A. (1997). On the importance of imagery in the cognitive interview. Reply to Roberts. *Psycholoquy*, *8*(3), Witness Memory (16). Retrieved October 20, 2006, from http://psycprints.ecs.soton.ac.uk/archive/00000539/

Stinson, V., Devenport, J.L., Cutler, B.L., & Kravitz, D.A. (1996). How effective is the presence-of-counsel safeguard? Attorney perceptions of suggestiveness, fairness and correctability of biased lineup procedures. *Journal of Applied Psychology*, *81*, 64–75.

Stinson, V., Devenport, J.L., Cutler, B.L., & Kravitz, D.A. (1997). How effective is the motion-to-suppress safeguard? Judges' perceptions of suggestiveness and fairness of biased lineup procedures. *Journal of Applied Psychology*, *82*, 211–220.

Tanaka, J.W., & Farah, M.J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A*, *46*, 225–246.

Tanaka, J.W., & Sengco, J.A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*, 583–592.

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). New York: Cambridge University Press.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 353–370.

Turtle, J.W., Lindsay, R.C.L., & Wells, G.L. (2003). Best practice recommendations for eyewitness evidence procedures: New ideas for the oldest way to solve a case. *Canadian Journal of Police and Security Services*, *1*, 5–18.

Valentine, T., & Heaton, P. (1999). An evaluation of the fairness of police line-ups and video identifications. *Applied Cognitive Psychology*, *13*, S59–S72.

Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, *17*, 969–993.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. Chichester, England: Wiley.

Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *Journal of Applied Psychology*, *88*, 490–499.

Weber, N., & Brewer, N. (2004). Confidence-accuracy calibration in absolute and relative face recognition judgments. *Journal of Experimental Psychology: Applied*, *10*, 156–172.

Weber, N., Brewer, N., Wells, G.L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: The unruly 10–12 second rule. *Journal of Experimental Psychology: Applied*, *10*, 139–147.

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

Wells, G.L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*, 89–103.

Wells, G.L. (1988). *Eyewitness identification: A system handbook*. Toronto: Carswell Legal Publications.

Wells, G.L. (1993). What do we know about eyewitness identification? *American Psychologist*, *48*, 553–571.

Wells, G.L. (2006). Eyewitness identification: Systemic reforms. *Wisconsin Law Review*, *2006*, 615–643.

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.

Wells, G.L., & Bradfield, A.L. (1999a). Distortions in eyewitnesses' recollections: Can the postidentification feedback effect be moderated? *Psychological Science*, *10*, 138–144.

Wells, G.L., & Bradfield, A.L. (1999b). Measuring the goodness of lineups: Parameter estimation, question effects, and limits to the mock witness paradigm. *Applied Cognitive Psychology*, *13*, 27–40.

Wells, G.L., Charman, S.D., & Olson, E.A. (2005). Building face composites can harm lineup identification performance. *Journal of Experimental Psychology: Applied*, *11*, 147–157.

Wells, G.L., Ferguson, T.J., & Lindsay, R.C.L. (1981). The tractability of eyewitness confidence and its implication for triers of fact. *Journal of Applied Psychology*, *66*, 688–696.

Wells, G.L., & Hryciw, B. (1984). Memory for faces: Encoding and retrieval operations. *Memory and Cognition*, *12*, 338–344.

Wells, G.L., & Leippe, M.R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Memory for peripheral detail can be misleading. *Journal of Applied Psychology*, *66*, 682–687.

Wells, G.L., Leippe, M.R., & Ostrom, T.M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior*, *3*, 285–293.

Wells, G.L., & Lindsay, R.C.L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, *88*, 776–784.

Wells, G.L., Lindsay, R.C.L., & Ferguson, T.J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology*, *64*, 440–448.

Wells, G.L., & Luus, E. (1990). Police lineups as experiments: Social methodology as a framework for properly conducted lineups. *Personality and Social Psychology Bulletin*, *16*, 106–117.

Wells, G.L., Malpass, R.S., Lindsay, R.C.L., Fisher, R.P., Turtle, J.W., & Fulero, S. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, *55*, 581–598.

Wells, G.L., & Murray, D.M. (1983). What can psychology say about the Neil vs. Biggers criteria for judging eyewitness identification accuracy? *Journal of Applied Psychology*, *68*, 347–362.

Wells, G.L., & Murray, D.M. (1984). Eyewitness confidence. In G.L. Wells & E.F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 155–170). New York: Cambridge University Press.

Wells, G.L., & Olson, E. (2002). Eyewitness identification: Information gain from incriminating and exonerating behaviors. *Journal of Experimental Psychology: Applied*, *8*, 155–167.

Wells, G.L., Olson, E., & Charman, S. (2002). Eyewitness identification confidence. *Current Directions in Psychological Science*, *11*, 151–154.

Wells, G.L., Olson, E., & Charman, S. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied*, *9*, 42–52.

Wells, G.L., Rydell, S.M., & Seelau, E.P. (1993). On the selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, *78*, 835–844.

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

Wells, G.L., & Turtle, J.W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin*, *99*, 320–329.

Westcott, H., Davies, G., & Bull, R. (2002). *Children's testimony: A handbook of psychological research and forensic practice*. Chichester, England: Wiley.

Whipple, G.M. (1909). The obtaining of information: Psychology of observation and report. *Psychological Bulletin*, *8*, 419–421.

Whipple, G.M. (1910). Recent literature on the psychology of testimony. *Psychological Bulletin*, *8*, 365–368.

Whipple, G.M. (1911). The psychology of testimony. *Psychological Bulletin*, *9*, 307–309.

Whipple, G.M. (1912). Psychology of testimony and report. *Psychological Bulletin*, *10*, 264–269.

Wigmore, J.H. (1909). Professor Münsterburg and the psychology of testimony. *Illinois Law Review*, *3*, 399–445.

Wogalter, M.S., Marwitz, D.B., & Leonard, D.C. (1992). Suggestiveness in photospread lineups: Similarity induces distinctiveness. *Applied Cognitive Psychology*, *6*, 443–453.

Wright, D.B., Boyd, C.E., & Tredoux, C.G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology*, *17*, 365–373.

Wright, D.B., & Loftus, E.F. (1998). How misinformation alters memories. *Journal of Experimental Child Psychology*, *71*, 155–164.

Wright, D.B., & McDaid, A.T. (1996). Comparing system and estimator variables using data from real lineups. *Applied Cognitive Psychology*, *10*, 75–84.

Wright, D.B., & Skagerberg, E.M. (in press). Post-identification feedback affects real eyewitnesses. *Psychological Science*.

Yarmey, A.D. (2004). Eyewitness recall and photo identification: A field experiment. *Psychology, Crime and Law*, *10*, 53–68.

Yuille, J., & Cutshall, J. (1986). A case study of eyewitnesses' memory of a crime. *Journal of Applied Psychology*, *71*, 291–301.

Zaragoza, M.S., Payment, K.E., Ackil, J.K., Drivdahl, S.B., & Beck, M. (2001). Interviewing witnesses: Forced confabulation and confirmatory feedback increase false memories. *Psychological Science*, *12*, 473–477.