Annu. Rev. Psychol. 2003. 54:277–95
doi: 10.1146/annurev.psych.54.101601.145028

First published online as a Review in Advance on August 27, 2002

# EYEWITNESS TESTIMONY

Gary L. Wells and Elizabeth A. Olson
*Psychology Department, Iowa State University, Ames, Iowa 50011;*
*e-mail: glwells@iastate.edu; vorndran@iastate.edu*

**Key Words** witness, testimony, lineup

■ **Abstract** The criminal justice system relies heavily on eyewitness identification for investigating and prosecuting crimes. Psychology has built the only scientific literature on eyewitness identification and has warned the justice system of problems with eyewitness identification evidence. Recent DNA exoneration cases have corroborated the warnings of eyewitness identification researchers by showing that mistaken eyewitness identification was the largest single factor contributing to the conviction of these innocent people. We review major developments in the experimental literature concerning the way that various factors relate to the accuracy of eyewitness identification. These factors include characteristics of the witness, characteristics of the witnessed event, characteristics of testimony, lineup content, lineup instructions, and methods of testing. Problems with the literature are noted with respect to both the relative paucity of theory and the scarcity of base-rate information from actual cases.

## CONTENTS



Eyewitnesses are critical in solving crimes, and sometimes eyewitness testimony is the only evidence available for determining the identity of the culprit. Psychological researchers who began programs in the 1970s, however, have consistently articulated concerns about the accuracy of eyewitness identification. Using various methodologies, such as filmed events and live staged crimes, eyewitness researchers have noted that mistaken identification rates can be surprisingly high and that eyewitnesses often express certainty when they mistakenly select someone from a lineup. Although their findings were quite compelling to the researchers themselves, it was not until the late 1990s that criminal justice personnel began taking the research seriously. This change in attitude about the psychological literature on eyewitness identification arose primarily from the development of forensic DNA tests in the 1990s. More than 100 people who were convicted prior to the advent of forensic DNA have now been exonerated by DNA tests, and more than 75% of these people were victims of mistaken eyewitness identification (Wells et al. 1998, Scheck et al. 2000). The apparent prescience of the psychological literature regarding problems with eyewitness identification has created a rising prominence of eyewitness identification research in the criminal justice system (Wells et al. 2000).

Because most crimes do not include DNA-rich biological traces, reliance on eyewitness identification for solving crimes has not been significantly diminished by the development of forensic DNA tests. Interestingly, research on eyewitness reliability has been done only by psychologists—primarily cognitive and social psychologists—and the psychological literature represents the only source of empirical data on eyewitness identification. The vast criminal justice system itself has never conducted an experiment on eyewitness identification.

## COVERAGE OF THIS REVIEW

No review of the eyewitness identification literature has previously appeared in the *Annual Review of Psychology*. Therefore, we include here references to articles from the 1970s and 1980s that we think especially critical to the development of the literature, but we primarily emphasize more recent developments. Also, because the eyewitness identification literature has become so vast, we are necessarily selective in our citations and coverage. Readers should note that this review focuses on eyewitness identification rather than on eyewitness testimony in general. Eyewitnesses commonly testify about many things, such as which hand a gunman used, the color of a car, or recollections of a conversation, but these event memories are outside the scope of this review. The large literature on child eyewitnesses, suggestibility, and recovery of repressed memories is not reviewed here.

## BASIC CONCEPTS

The eyewitness identification literature has developed a number of definitions and concepts that require explanation. A *lineup* is a procedure in which a criminal suspect (or a picture of the suspect) is placed among other people (or pictures of other people) and shown to an eyewitness to see if the witness will identify the suspect as the culprit in question. The term suspect should not be confused with the term culprit. A *suspect* might or might not be the *culprit* (a suspect is suspected of being the culprit). *Fillers* are people in the lineup who are not suspects. Fillers, sometimes called *foils* or *distractors*, are known-innocent members of the lineup. Therefore, the identification of a filler would not result in charges being brought against the filler. A *culprit-absent* lineup is one in which an innocent suspect is embedded among fillers and a *culprit-present* lineup is one in which a guilty suspect (culprit) is embedded among fillers. The primary literature sometimes calls these *target-present* and *target-absent* lineups.

A *simultaneous* lineup is one in which all lineup members are presented to the eyewitness at once and is the most common lineup procedure in use by law enforcement. A *sequential* lineup, on the other hand, is one in which the witness is shown only one person at a time but with the expectation that there are several lineup members to be shown.

A lineup's *functional size* is the number of lineup members who are "viable" choices for the eyewitness. For example, if the eyewitness described the culprit as being a tall male with dark hair and the suspect is the only lineup member who is tall with dark hair, then the lineup's functional size would be 1.0 even if there were 10 fillers. Functional size was introduced as a specific measure (Wells et al. 1979), and competing measures have been proposed, such as Malpass's (1981) "effective size." Today functional size is used generically to mean the number of lineup members who fit the eyewitness's description of the culprit.

*Mock witnesses* are people who did not actually witness the crime but are asked to pick a person from the lineup based on the eyewitness's verbal description of the culprit. Mock witnesses are used to test the functional size of the lineup.

The *diagnosticity* of suspect identification is the ratio of accurate identification rate with a culprit-present lineup to the inaccurate identification rate with a culprit-absent lineup. The *diagnosticity of "not there"* is the ratio of "not there" response rates with culprit-absent lineups to "not there" response rates with culprit-present lineups. The *diagnosticity of filler identifications* is the ratio of filler identification rates with culprit-absent lineups to filler identification rates with culprit-present lineups.

Among variables that affect eyewitness identification accuracy, a *system variable* is one that is (or could be) under control of the criminal justice system, while an *estimator variable* is one that is not. System variables include instructions given to eyewitnesses prior to viewing a lineup and the functional size of a lineup. Estimator variables include lighting conditions at the time of witnessing and whether the witness and culprit are of the same or of different races.

The distinction between estimator and system variables has assumed great significance in the eyewitness identification literature since it was introduced in the late 1970s (Wells 1978). In large part, the prominence of this distinction attests to the applied nature of the eyewitness identification literature. Whereas the development of a literature on estimator variables permits some degree of postdiction that might be useful for assessing the chances of mistaken identification after the fact, the development of a system variable literature permits specification of how eyewitness identification errors might be prevented in the first place.

## ESTIMATOR VARIABLES

Estimator variables can be sorted into four broad categories: characteristics of the witness, characteristics of the event, characteristics of the testimony, and abilities of the testimony evaluators to discriminate between accurate and inaccurate witness testimony.

### Characteristics of the Witness

Are members of certain groups better eyewitnesses than those of others? The empirical evidence is not overwhelming. For example, there is no clear evidence that males and females differ significantly overall in ability to identify people from lineups. A meta-analysis by Shapiro & Penrod (1986) indicated that females might be slightly more likely to make accurate identifications but also slightly more likely to make mistaken identifications than are males (due to females being more likely to attemp an identification), thereby yielding an overall equivalent diagnosticity for males and females. Although males and females might take an interest in different aspects of a scene and thereby remember somewhat different details (e.g., Powers et al. 1979), overall abilities of males and females in eyewitness identification appear to be largely indistinguishable (but see Brigham & Barkowitz 1978, Shaw & Skolnick 1999).

The age of the eyewitness, on the other hand, has been consistently linked to eyewitness identification performance, with very young children and the elderly performing significantly worse than younger adults. The eyewitness identification errors of young children and the elderly are highly patterned: When the lineup contains the actual culprit, young children and the elderly perform nearly as well as young adults in identifying the culprit, but when the lineup does not contain the culprit the young children and the elderly commit mistaken identifications at a higher rate than do young adults (see the meta-analysis on children versus adults by Pozzulo & Lindsay 1998).

There is little evidence that intelligence is related to eyewitness identification performance. Although an early study by Howells (1938) indicated a significant relation between face recognition accuracy and intelligence, later studies have shown no relation (e.g., Brown et al. 1977). A word of caution is in order here, however, because Howells's sample of witnesses included a much greater range of intelligence at the low end than have later studies. At the low extremes of intelligence, a pattern similar to that found with children seems likely, namely a high rate of mistaken identifications in response to culprit-absent lineups.

The race of the eyewitness has been examined extensively. Although no consistent overall differences attributable to race have emerged, the evidence is now quite clear that people are better able to recognize faces of their own race or ethnic group than faces of another race or ethnic group. A recent meta-analysis by Meissner & Brigham (2001) shows that this effect is robust across more than 25 years of research.

Little published research relates personality characteristics to eyewitness identification accuracy. Hosch et al. (1984) found that high self-monitors (individuals who adapt their behavior to cues regarding what is socially appropriate) are more susceptible to biased lineup procedures than are low self-monitors, and Hosch & Platz (1984) found a relation between self-monitoring and correct identifications. Also, a meta-analysis by Shapiro & Penrod (1986) indicated that individuals high in chronic trait anxiety (a general attitude of apprehension) made fewer mistaken identifications than individuals low in chronic trait anxiety. Their meta-analysis also indicated that field independents (those with a perceptual tendency to differentiate parts of a visual field from the whole) made fewer accurate identifications (but equal mistaken identifications) than did field dependents. However, little research has been directed at the role of personality in eyewitness identification, and no strong theory relating personality to eyewitness identification has emerged.

## Characteristics of the Event

A variety of factors affect the ability of an eyewitness to identify the culprit at a later time, including the amount of time the culprit is in view, the lighting conditions, whether the culprit wears a disguise, the distinctiveness of the culprit's appearance, the presence or absence of a weapon, and the timing of knowledge that one is witnessing a crime.

Distinctive faces are much more likely to be accurately recognized than nondistinctive faces (e.g., Light et al. 1979). Faces that are highly attractive or highly unattractive are easier to recognize than are faces that are average in attractiveness (e.g., Fleishman et al. 1976), but what makes a face distinctive is not entirely clear. Because the arithmetic mean (averaged at the pixel level) of several faces (a prototype) is judged to be more attractive than the individual faces that were averaged (see Langois & Roggman 1990), the distinctiveness-recognition relation is probably not due to a simple deviation from the arithmetic mean of individual facial features.

Simple disguises, even those as minor as covering the hair, result in significant impairment of eyewitness identification (Cutler et al. 1987). Sunglasses also impair identification, although the degree of impairment can be reduced by having the targets wear sunglasses at the time of the recognition test (Hockley et al. 1999). Photos of criminal suspects used in police lineups are sometimes several years old. Changes in appearance that occur naturally over time and changes that are made intentionally by suspects can have quite strong effects on recognition. Read et al. (1990) found that photos of the same people taken two years apart were less likely to be recognized as the same people when their appearance had naturally changed (via aging, facial hair) than when their appearance had remained largely the same.

Clearly, at the extreme of low light levels there is a point at which a face cannot be perceived well enough to be recognized later. Surprisingly, however, we know of no experiments that have measured the light levels required for the encoding of faces. We encourage researchers to address this question.

As would be expected, the amount of time a culprit's face is in view affects the chances that the eyewitness can identify the person later (Ellis et al. 1977). However, this relationship depends less critically on the eyewitness's opportunity to view per se and more on the amount and type of attention that the witness directs at the culprit. Given equal exposure time to a face, people are more likely to be able to recognize that face later if they make abstract inferences about it (e.g., is this person honest?) than if they make physical judgments (e.g., does this person have a large or small nose?). Presumably, this effect occurs because the abstract inferences require holistic processing of the face whereas the physical judgments require feature processing (Wells & Hryciw 1984).

In general, the amount of time a culprit's face is in view is not as critical for eyewitness identification accuracy as the type or amount of attention given by the witness. For example, Leippe et al. (1978) exposed unsuspecting people to a staged theft of a package. Some were led to believe that the package contained a valuable item and some were led to believe that the package contained a trivial item. In addition, some learned of the value of the item in the package before the theft and some only learned the value after the thief had fled. Although all had the same opportunity to view the thief, the witnesses who knew the value of the item beforehand were significantly more accurate at identification than the other three groups. Observers often do not realize that they have witnessed a crime until after the culprit has fled. Although they might have had significant opportunity to view the culprit, they might have had little reason to attend closely.

One factor that can signal to eyewitnesses that a crime is occurring is the presence of a weapon. Unfortunately, learning that one is an eyewitness to a crime via the culprit's display of a weapon might not make the person a better eyewitness. A number of studies have been directed at the question of the so-called weapon-focus effect. A meta-analysis of these studies indicates that the presence of a weapon reduces the chances that the eyewitness can identify the holder of the weapon (Steblay 1992). Loftus et al. (1987) monitored eyewitnesses's eye movements and found that weapons draw visual attention away from other things such as the culprit's face. Complicating the issue somewhat is the fact that the presence of weapons or other types of threatening stimuli can cause arousal, fear, and emotional stress. The effects of such stress on memory are still being debated. Some research shows that increased levels of violence in filmed events reduces eyewitness identification accuracy (e.g., Clifford & Hollin 1981) whereas other research has failed to find this effect (e.g., Cutler et al. 1987). Deffenbacher (1983) suggested that the effect is likely to follow the Yerkes-Dodson Law where only very high and very low levels of arousal will impair memory. Christianson's (1992) review of the evidence relating emotional stress to memory suggests that emotional events receive preferential processing; emotional response causes a narrowing of attention (as suggested by Easterbrook 1959) with loss of peripheral details.

## Characteristics of Testimony

Considerable interest and research have been directed at the question of whether there are characteristics of an eyewitness's testimony that could be used to postdict whether the witness made an accurate or false identification. The bulk of this research has focused on the certainty (confidence) of the eyewitness. Although early research suggested that the certainty an eyewitness expresses in an identification is largely unrelated to the accuracy of the identification, current analyses suggest a more hopeful but also more complex view of the certainty-accuracy relation. Although any given experiment might show a statistically nonsignificant relation between certainty and accuracy, meta-analyses of the literature show a reliable correlation. Several moderators of the strength of the relation have been identified. One important moderator is the overall accuracy of the eyewitnesses. When accuracy is low (e.g., from poor witnessing conditions), the certainty-accuracy relationship suffers (Bothwell et al. 1987). Later meta-analyses indicate that the certainty-accuracy relation is stronger if the analysis is restricted to those making an identification (choosers only) than if it also includes witnesses who make correct and false rejections (Sporer et al. 1995). In fact, using a weighted average of effect sizes for choosers only, Sporer et al. reported a 0.37 certainty-accuracy correlation across 30 studies. More recent work indicates that directing eyewitnesses to reflect on their encoding and test conditions or asking them to entertain hypotheses regarding why their identification might have been mistaken can improve the relation between accuracy and certainty, especially when this relation is calculated using calibration methods rather than the point-biserial correlation (Brewer et al. 2002).

Although the 0.37 correlation estimate for the certainty-accuracy relation is more optimistic than the early estimates, recent studies suggest the literature might be overestimating the utility of eyewitness certainty in actual cases. In a series of experiments, eyewitness certainty was shown to be highly malleable among eyewitnesses who had made mistaken identifications (Wells & Bradfield 1998, 1999). After making mistaken identifications, some eyewitnesses were given confirming feedback by the lineup administrator ("Good, you identified the suspect") whereas others were given no feedback about their identification. This feedback served to distort the eyewitnesses' recollections of the certainty they had in their identifications. Those given confirming feedback recalled having been very certain in their identification compared to those given no confirming feedback. This certainty-inflation effect is greater for eyewitnesses who make mistaken identifications than it is for those who make accurate identifications, resulting in a significant loss in the certainty-accuracy relation (Bradfield et al. 2002). In actual cases, it is common for lineup administrators (usually the detective in the case) to give confirming feedback to eyewitnesses, thereby inflating the certainty of the eyewitness and confounding the certainty-accuracy relation. Even if the lineup administrator refrains from giving the witness confirming feedback, the witness is likely to make confirming inferences from later events (e.g., an indictment of the identified person). Another real-world factor that can muddle the meaning of eyewitness certainty is repeated testing. Shaw and his colleagues (Shaw 1996, Shaw & McClure 1996) have shown that repeated questioning of eyewitnesses on a matter about which they were inaccurate serves to inflate their certainty that they were accurate. Hence, it is unclear whether the .37 correlation between certainty and accuracy revealed in the Sporer et al. meta-analysis of experiments can be directly applied to actual cases in which there are other influences that inflate the certainty of eyewitnesses.

An even more promising indicator of eyewitness accuracy is the speed with which the eyewitness makes an identification from a lineup. Several studies have now found that witnesses who make accurate identifications from a lineup reach their decision faster than do witnesses who make mistaken identifications (Dunning & Perretta 2002; Dunning & Stern 1994; Robinson et al. 1997; Smith et al. 2000; Sporer 1992, 1993, 1994). In an impressive set of results, Dunning & Perretta found that those who made their decision in less than 10–12 seconds were nearly 90% accurate in their identifications from a lineup whereas those taking longer were approximately 50% correct. The 10–12-second rule was developed post hoc to produce the best separation of accurate and inaccurate witnesses, so some caution is called for with regard to how well the 10–12-second rule works in other situations; but the general relation between accuracy and speed of identification has received support in several studies. In addition, the idea that faster identifications are more likely to be accurate than are slower identifications makes good theoretical sense. It has long been theorized that mistaken identifications result from a deliberated judgment in which witnesses compare one lineup member to another and use inferences and elimination strategies to decide which person must be the culprit whereas accurate identifications result from a more automatic recognition process that does not require comparisons of one lineup member to another (Wells 1984a).

## Lay Observers' Judgments of Accuracy

Observers (e.g., jurors) have little ability to make correct discriminations between accurate and inaccurate eyewitness identification testimony. Several methods have been used to assess the adequacy of people's judgments about eyewitness identification accuracy. Surveys, for example, show poor agreement (often less than 50%) between the answers that lay people give about variables affecting eyewitness identification accuracy and the answers researchers score correct based on the empirical literature (e.g., Deffenbacher & Loftus 1982, McConkey & Roche 1989, Noon & Hollin 1987). Another approach has been to use "prediction" studies in which eyewitness identification experiments are described and people are asked to predict the results. The results of these studies show a tendency to overestimate eyewitness identification accuracy and a failure to correctly predict interactions between variables (e.g., Brigham & Bothwell 1983, Wells 1984b).

A third approach is to cross-examine eyewitnesses to staged crimes and to ask subject-jurors to determine whether witnesses made accurate or mistaken identifications. In a series of experiments using this methodology, subject-jurors have shown little or no ability to make such discriminations (Lindsay et al. 1989, Lindsay et al. 1981, Wells et al. 1981, Wells & Leippe 1981, Wells et al. 1979). Because observers' belief rates exceeded eyewitnesses' accuracy rates, these studies are commonly cited as evidence that people are overbelieving of eyewitnesses. However, this pattern of overbelief is restricted primarily to poorer witnessing conditions; when witnessing conditions were good, belief rates and eyewitness identification accuracy rates were more similar. In addition, mock jurors sometimes underbelieved the eyewitnesses who had quite low levels of certainty.

## SYSTEM VARIABLES

System variables are those that affect the accuracy of eyewitness identifications and over which the criminal justice system has (or can have) control. In general, these tend to be lineup test factors, such as how witnesses are instructed prior to viewing a lineup or how the lineup is structured. The distinction between system variables and estimator variables is consequential in several respects. Whereas estimator variables can at best increase the probability that the criminal justice system can sort accurate from inaccurate eyewitness identifications, system variables can help prevent inaccurate identifications from occurring in the first place. Consider, for instance, the idea that jurors tend to overbelieve eyewitness identification testimony. Although expert testimony about eyewitness identification might manage to reduce jurors' tendencies to overestimate eyewitnesses' accuracy, the system variable approach might enable eyewitness identification accuracy to match the level of jurors' beliefs (Seelau & Wells 1995).

The procedure used by crime investigators conducting a lineup has been likened to that of researchers conducting an experiment (Wells & Luus 1990). Crime investigators begin with a hypothesis (that the suspect is the culprit), create a design for testing the hypothesis (embed the suspect among fillers), carry out a procedure (e.g., provide pre-lineup instructions and present the group to an eyewitness), observe and record the eyewitness's behavior (witness decision), and then interpret and revise their hypothesis (whether the suspect is the culprit). All the types of things that can go wrong with an experiment to cause misleading results can also go wrong with a lineup. For instance, the instructions might bias the witness, the hypothesis might be prematurely leaked, the design might be flawed, the behavior might be misinterpreted, confirmation biases might be operating, and so on. Indeed, a great deal of the research literature on system variable eyewitness identification could be construed as the extension of sound experimental methodology to the design and procedure of police lineups.

Most system variable research in eyewitness identification can be placed into four categories: instructions, content, presentation method, and behavioral influence. Before reviewing these system variables, however, it is important to understand the role played by the presence versus absence of the culprit in the lineup and the concept of a relative-judgment decision process.

## Culprit-present Versus Culprit-absent Lineups

A lineup might or might not include the actual culprit. If police investigators have unknowingly focused on an innocent person as their suspect and place that suspect in the lineup, then the eyewitness(es) will end up viewing a lineup for which the only correct answer is "not there." Research repeatedly shows that culprit-absent lineups present great problems for eyewitnesses. The same eyewitnesses who identified an innocent person from a culprit-absent lineup might otherwise have been able to identify the actual culprit from a culprit-present lineup (Wells 1984a). In one study, for example, 54% of eyewitnesses were able to identify the actual culprit from a 6-person culprit-present lineup and 21% made no identification. When the culprit was removed without replacement (making it a 5-person culprit-absent lineup), however, the rate of no identification rose only to 32%, with the other 68% of the eyewitnesses who saw this lineup mistakenly identifying someone from the 5 remaining members of the lineup (Wells 1993).

A theoretical view that has been used heavily in the eyewitness identification literature is that eyewitnesses tend to use a relative-judgment decision process in making identifications from a lineup (Wells 1984a). The relative-judgment conceptualization states that an eyewitness tends to select a person from a lineup who most resembles the eyewitness's memory of the culprit relative to the other lineup members. Although the relative-judgment decision process permits eyewitnesses to do a reasonable job of identifying the culprit from a culprit-present lineup, when eyewitnesses view a culprit-absent lineup there will likely be one lineup member who looks more like the culprit than the others.

An alternative explanation of the errors witnesses make with culprit-absent lineups is that eyewitnesses tend to have lax criteria of resemblance; under culprit-absent circumstances, innocent lineup members easily meet these undemanding criteria (Ebbesen & Flowe 2002). Experimental data have not yet favored one of these interpretations over the other. Recent mathematical modeling of lineup data by Clark may help to refine our understanding of the roles of both relative judgments and criterion setting (Clark 2002).

## Instructions

A variable shown repeatedly to have considerable impact on eyewitness identifications from lineups is the pre-lineup instruction given to eyewitnesses. Malpass & Devine (1981) were the first to demonstrate that the ratio of accurate to inaccurate identifications is strongly affected by whether or not eyewitnesses have been instructed (warned) prior to viewing the lineup that the culprit might or might not be in the lineup. A meta-analysis of the eyewitness identification literature on pre-lineup instructions reveals that the loss of accurate identifications from such instructions is minimal whereas the reduction of mistaken identifications is considerable (Steblay 1997). Steblay's meta-analysis showed that the presence of the "might or might not be present" instruction (compared to no instruction) reduced mistaken identification rates in culprit-absent lineups by 41.6% whereas accurate identification rates in culprit-present lineups were reduced by only 1.9%. Based on this compelling research, the U.S. Department of Justice included this type of instruction in its first set of national guidelines for law enforcement on the collection of eyewitness evidence (Technical Working Group for Eyewitness Evidence 1999).

## Lineup Content

When police have a suspect and decide to conduct a lineup, nonsuspect (filler) members of the lineup must be chosen. The importance of the selection of fillers as a system variable was demonstrated early, and it remains one of the primary active issues in the eyewitness identification literature. Ideally, lineup fillers would be chosen so that an innocent suspect is not mistakenly identified merely from "standing out," and so that a culprit does not escape identification merely from blending in. The first experimental demonstration of the importance of filler selection showed what can happen when this idea is not achieved. When fillers did not at all resemble the culprit, eyewitnesses tended to mistakenly identify an innocent suspect who resembled the culprit; when the suspect was the culprit, however, the manipulation of fillers had little effect on the rate of accurate identifications (Lindsay & Wells 1980).

Although the issue of lineup fillers seems simple at first glance, it is in fact complex. In the early demonstrations, researchers used their knowledge of the culprit's identity to select fillers. In actual cases, of course, the identity of the culprit is not known. Using the suspect as a proxy for the culprit will have different effects on rates of accurate and mistaken identification depending on whether the suspect is the culprit or an innocent person. Accordingly, selecting fillers who are highly similar to the suspect can help protect the innocent suspect in a culprit-absent lineup, but can also reduce accurate identifications in a culprit-present lineup (Luus & Wells 1991). Another line of research has shown that using the suspect as the reference point to select fillers can create a "backfire effect" in which an innocent suspect, being the origin or central tendency of the lineup, actually has an increased chance of being identified as the culprit (Clark & Tunnicliff 2001, Navon 1992, Wogalter et al. 1992).

An alternative to the strategy of selecting fillers based on their resemblance to the suspect is to select fillers based on their fit to the verbal description the eyewitness had given of the culprit. This fit-to-description strategy has several practical advantages (see Wells et al. 1994) and has worked well in some experiments (Juslin et al. 1996, Wells et al. 1993). However, biases against the innocent suspect can remain with the fit-to-description method when the description is especially sparse or when the innocent suspect happens to show a high resemblance to the culprit (Clark & Tunnicliff 2001, Lindsay et al. 1994). In actual cases, high resemblance between the innocent suspect and the culprit can occur by chance or it can occur whenever the innocent person became a suspect because she or he resembled a composite or a security video image of the culprit.

## Lineup Presentation Method

Many alternatives to the traditional lineup have been proposed and tested, and future research will likely focus on solving the lineup system variable problems. The first proposed alternative to the traditional lineup was the blank lineup control method (Wells 1984a). A blank lineup is one that contains only fillers (no suspect). The eyewitness is first shown the blank lineup under the belief that this is the only lineup to be shown. The identification of someone from a blank lineup is known to be an error (because the lineup members are all fillers), and witnesses who make an identification from a blank lineup can thereby be discarded. Witnesses who do not make an identification from the blank lineup can then be shown the actual lineup (which contains a suspect). Data indicate that eyewitnesses who do not make an identification from the blank lineup are much more reliable on the second (actual) lineup than are those who were not screened with the blank lineup method. In effect, the blank lineup method is analogous to the use of a control condition in a within-subjects design and could be used in actual cases. In general, however, crime investigators have not liked the idea of the blank lineup control method on grounds that it "tricks" the eyewitness and could sever the eyewitness's trust in investigators.

Another proposed alternative to the traditional lineup procedure, and the best known of these alternatives, is the sequential lineup (Lindsay & Wells 1985). Unlike the traditional lineup in which the lineup members are shown to the eyewitness simultaneously, the sequential lineup shows the eyewitness only one lineup member at a time and requires the eyewitness to make a decision ("Is this person the culprit or not?") prior to viewing the next lineup member. The most powerful version of the sequential procedure is one in which the eyewitness does not know how many lineup members are to be viewed. In theory, the sequential lineup procedure prevents eyewitnesses from selecting the person who looks most like the culprit relative to the other lineup members, a process called relative-judgment decision (see above) (Wells 1984a). To the extent that relative judgments are operating, eyewitnesses will have difficulty with culprit-absent lineups because by definition someone in the lineup resembles the culprit more closely than the other lineup members do. Unlike the simultaneous lineup, the sequential lineup prevents eyewitnesses from making a relative-judgment decision because at any point in the sequence a lineup member who has not yet been viewed may turn out to resemble the culprit more than any person viewed thus far. Eyewitnesses must compare each member of the sequential lineup to their memory of the culprit and thus make a more "absolute judgment" about identity. A recent meta-analysis of 25 studies comparing simultaneous and sequential lineups showed that the sequential lineup reduced the chances of mistaken identifications in culprit-absent lineups by nearly one half (Steblay et al. 2001). Unfortunately, the sequential technique was also associated with a reduction in accurate identification rates in culprit-present lineups. Although this reduction was not as great as that in mistaken identifications, it was nevertheless statistically reliable. The pattern of these results has led Ebbesen &

Flowe (2002) to speculate that the sequential lineup raises the criteria for making a positive identification rather than changing the process from relative to absolute judgments.

Another alternative to the traditional lineup is the elimination lineup, a procedure in which the witness's task is to eliminate all but one lineup member and then make a separate decision as to whether that person is the culprit or not (Pozzulo & Lindsay 1999). Although the elimination lineup does not seem to work well with adults, it seems to eliminate some of the problems young children have with lineups.

## Behavioral Influence: The Need for Double-Blind Testing

One of the ways that the justice system itself can influence eyewitness identification evidence is through the behaviors of the person who administers the lineup (Wells 1993). Commonly, the person who administers a lineup is the case detective who, of course, knows which member of the lineup is the suspect and which members are fillers. The need for double-blind testing is well established in the behavioral sciences (Rosenthal 1976) but is largely unknown or unheeded in criminal investigation procedures and forensic science (Risinger et al. 2002). Lineup administrators could inadvertently communicate their knowledge about which lineup member is the suspect and which members are merely fillers to the eyewitness through various verbal and nonverbal means. Phillips et al. (1999) manipulated lineup administrators' assumptions about the identity of the culprit and found that this manipulation affected the choices that eyewitnesses made from the lineup, especially when a sequential lineup procedure was used. In addition to influencing eyewitnesses' choice of particular lineup members, the person administering the lineup can cause other problems. Wells & Bradfield (1998, 1999) found that post-identification suggestions to eyewitnesses from lineup administrators led mistaken eyewitnesses to develop high levels of false certainty that they had made an accurate identification. The problem of influence from the lineup administrator is easily fixed by having lineups administered by someone who does not know which lineup member is the suspect and which ones are fillers (Wells et al. 1998).

## Base Rates as System Variables

Base rates can be considered system variables in some cases. The important base rate in eyewitness identification is the base rate for the culprit being present versus absent in a lineup. Most mistaken identifications occur when the culprit is not in the lineup. Although the relation between the culprit-absent versus culprit-present base rate and the chances of mistaken identification has been established and modeled mathematically (Wells & Lindsay 1980, Wells & Turtle 1986), the case was only recently made for treating this base rate as a system variable (Wells & Olson 2002). Previously, this base rate was treated as a fixed (albeit largely unknown) variable in actual cases. In fact, however, no laws or rules determine when a suspect is placed in a lineup and, therefore, this base rate varies as a function of the decisions crime investigators make when conducting a lineup. Consider, for instance, two police departments, a lax-criterion department and a strict-criterion department. In the lax-criterion department, investigators will place a suspect in a lineup for the slightest of reasons (e.g., a mere hunch) whereas the strict-criterion department requires certain evidence against a person (e.g., possession of stolen goods) before placing that person in a lineup. These two departments will, over the long run, have different base rates for culprit-present and culprit-absent lineups. Suppose, for example, that over a run of 1000 lineups the lax-criterion department shows 500 culprit-absent lineups and 500 culprit-present lineups whereas the strict-criterion department shows only 100 culprit-absent lineups and 900 culprit-present lineups. Given equivalent eyewitnesses in both of these departments, mistaken identifications of suspects will be nine times as likely in the lax-criterion department than in the strict-criterion department. (These surprising differences in the chances that an identification will be mistaken are simple derivations from Bayes' theorem.) Although the justice system has not yet done so, it could control the culprit-present versus culprit-absent base rate by requiring "probable cause" before placing someone in a lineup (Wells & Olson 2002). The base rate for culprit-present and culprit-absent lineups might be the most powerful system variable affecting the chances of mistaken identification.

## PROBLEMS AND PROSPECTS

In spite of the successful application of the eyewitness identification literature, significant work has yet to be done. The eyewitness identification literature has been driven much less by theoretical frameworks than by practical perspectives. Two problems are related to this state of affairs. One is that the premium on application and forensic relevance reduces the interplay and sharing of ideas between eyewitness identification researchers and their counterparts in basic areas of psychology, especially cognitive and social psychology. In addition, the experimental eyewitness identification literature is likely never to be complete enough to cover every possible situation that arises in actual cases; hence, better theory is needed to generalize this body of literature and to fill in gaps regarding what is likely to happen under various conditions.

A second concern is that while laboratory data on eyewitness identification are extensive, some key forms of real-world data are lacking. Certain estimable rates of eyewitness identification behavior and lineup conditions from actual cases could assist the design and interpretation of laboratory work. For instance, there have been no empirical estimates of the base rate for culprit-present versus culprit-absent lineups in actual cases. Although it is difficult to establish the ground truth (actual guilt or innocence) needed for precise estimates of this base rate in actual cases, methods exist for estimating upper limits (see Wells & Olson 2002). In addition, although the identification of a suspect from a lineup usually cannot be definitively classified as an accurate or mistaken identification in an actual case, the identification of a filler is a known error in actual cases and the rate at which these known errors occur can be informative. Two estimates of the filler identification rates in actual cases have been published. Wright & McDaid (1996) reported a rate of about 20% and Behrman & Davey (2001) reported a rate of 24%. One problem in collecting filler identification data from real cases is that police records often do not distinguish between eyewitnesses who make identifications of a filler and those who make no identification, which can result in a serious underestimation of the rate of filler identifications (Tollestrup et al. 1994). Another problem is that filler identification records from actual cases often lack an indication of the level of eyewitness certainty. These problems can be avoided by scripting data collection with police departments.

Actual case data of these types (e.g., base rates, filler identification rates, eyewitness certainty on known errors) can supplement the laboratory literature on eyewitness identification in two important ways. First, actual case data can be compared to laboratory data to see if the general rates of certain behaviors (e.g., nonidentification responses) are similar. Second, the rates for certain conditions in actual cases (e.g., rates of culprit-present versus culprit-absent lineups) are critical for Bayesian estimations of posterior probabilities that cannot themselves be derived from experiments.

Eyewitness identification research is likely to continue to focus on system variables for the foreseeable future because of the way system variables can be mapped onto the problem of improving eyewitness identification accuracy in actual cases. At the same time, estimator variables might be re-emerging with new promise for postdiction for three reasons. First, conditions are being found in which eyewitness certainty might be more closely related to eyewitness identification accuracy than once thought, especially when external influences on eyewitness certainty are minimized. Second, new postdiction variables, such as decision time, are emerging. Third, Bayesian analyses are being used to show that some eyewitness responses to lineups, such as filler identifications, have postdiction value in exonerations. Each of these represent potentially superior estimator variables because they can be more precisely measured in actual cases than can some of the more traditional estimator variables (such as stress or arousal). In any case, there is little evidence that eyewitness identification research is veering away from its applied orientation, especially in the face of recent successes in affecting legal policies and practices (Wells 2001, Wells et al. 2000).

**The *Annual Review of Psychology* is online at http://psych.annualreviews.org**

## LITERATURE CITED

Behrman BW, Davey SL. 2001. Eyewitness identification in actual criminal cases: an archival analysis. *Law Hum. Behav.* 25:475–91

Bothwell RK, Deffenbacher KA, Brigham JC. 1987. Correlation of eyewitness accuracy and confidence: optimality hypothesis revisited. *J. Appl. Psychol.* 72:691–95

Bradfield AL, Wells GL, Olson EA. 2002. The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *J. Appl. Psychol.* 87: 112–20

Brewer N, Keast A, Rishworth A. 2002. Improving the confidence-accuracy relation in eyewitness identification: evidence from correlation and calibration. *J. Exp. Psychol.: Appl.* 8:44–56

Brigham JC, Barkowitz P. 1978. Do "They all look alike?" The effect of race, sex, experience, and attitudes on the ability to recognize faces. *J. Appl. Soc. Psychol.* 8:306–18

Brigham JC, Bothwell RK. 1983. The ability of prospective jurors to estimate the accuracy of eyewitness identifications. *Law Hum. Behav.* 7:19–30

Brown E, Deffenbacher K, Sturgill W. 1977. Memory for faces and the circumstances of the encounter. *J. Appl. Psychol.* 62:311–18

Christianson SA. 1992. Emotional stress and eyewitness memory: a critical review. *Psychol. Bull.* 112:284–309

Clark SE. 2002. A memory and decision model for eyewitness identification. *Appl. Cogn. Psychol.* In press

Clark SE, Tunnicliff JL. 2001. Selecting lineup foils in eyewitness identification: experimental control and real-world simulation. *Law Hum. Behav.* 25:199–216

Clifford BR, Hollin CR. 1981. Effects of the type of incident and the number of perpetrators on eyewitness memory. *J. Appl. Psychol.* 66:364–70

Cutler BL, Penrod SD, Martens TK. 1987. The reliability of eyewitness identification: the role of system and estimator variables. *Law Hum. Behav.* 11:233–58

Deffenbacher K. 1983. The influence of arousal on reliability of testimony. In *Evaluating Witness Evidence: Recent Psychological Research and New Perspectives*, ed. SMA Lloyd-Bostock, BR Clifford, pp. 235–51. Chichester, Engl.: Wiley

Deffenbacher K, Loftus EF. 1982. Do jurors share a common understanding concerning eyewitness behavior? *Law Hum. Behav.* 6:15–30

Dunning D, Perretta S. 2002. Automaticity and eyewitness accuracy: a 10- to-12 second rule for distinguishing accurate from inaccurate positive identifications. *J. Appl. Psychol.* In press

Dunning D, Stern LB. 1994. Distinguishing accurate from inaccurate identifications via inquiries about decision processes. *J. Pers. Soc. Psychol.* 67:818–35

Easterbrook JA. 1959. The effect of emotion on cue utilization and the organization of behavior. *Psychol. Rev.* 66:183–201

Ebbesen E, Flowe S. 2002. Simultaneous versus sequential lineups: what do we really know? *Law Hum. Behav.* In press

Ellis HD, Davies GM, Shepherd JW. 1977. Experimental studies of face identification. *J. Crim. Def.* 3:219–34

Fleishman JJ, Buckley ML, Klosinsky MJ, Smith N, Tuck B. 1976. Judged attractiveness in recognition memory of women's faces. *Percept. Mot. Skills* 43:709–10

Hockley WE, Hemsworth DH, Consoli A. 1999. Shades of the mirror effect: recognition of faces with and without sunglasses. *Mem. Cogn.* 27(1):128–38

Hosch HM, Leippe MR, Marchioni PM, Cooper DS. 1984. Victimization, self-monitoring, and eyewitness identification. *J. Appl. Psychol.* 69:280–88

Hosch HM, Platz SJ. 1984. Self-monitoring and eyewitness accuracy. *Personal. Soc. Psychol. Bull.* 10:283–89

Howells TH. 1938. A study of ability to recognize faces. *J. Abnorm. Psychol.* 33:124–27

Juslin P, Olson 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. *J. Exp. Psychol. Learn. Mem. Cogn.* 5:1304–16

Langois JH, Roggman LA. 1990. Attractive faces are only average. *Psychol. Sci.* 1:115–21

Leippe MR, Wells GL, Ostrom TM. 1978. Crime seriousness as a determinant of accuracy in eyewitness identification. *J. Appl. Psychol.* 63:345–51

Light LL, Kayra-Stuart F, Hollander S. 1979. Recognition memory for typical and unusual faces. *J. Exp. Psychol.: Hum. Learn.* 5:212–28

Lindsay RCL, Martin R, Webber L. 1994. Default values in eyewitness descriptions: a problem for the match-to-description lineup foil selection strategy. *Law Hum. Behav.* 18: 527–41

Lindsay RCL, Wells GL. 1980. What price justice? Exploring the relationship between lineup fairness and identification accuracy. *Law Hum. Behav.* 4:303–14

Lindsay RCL, Wells GL. 1985. Improving eyewitness identification from lineups: simultaneous versus sequential lineup presentations. *J. Appl. Psychol.* 70:556–64

Lindsay RCL, Wells GL, O'Connor F. 1989. Mock juror belief of accurate and inaccurate eyewitnesses: a replication. *Law Hum. Behav.* 13:333–40

Lindsay RCL, Wells GL, Rumpel C. 1981. Can people detect eyewitness identification accuracy within and between situations? *J. Appl. Psychol.* 66:79–89

Loftus EF, Loftus GR, Messo J. 1987. Some facts about "weapon focus." *Law Hum. Behav.* 11:55–62

Luus CAE, Wells GL. 1991. Eyewitness identification and the selection of distracters for lineups. *Law Hum. Behav.* 15:43–57

Malpass RS. 1981. Effective size and defendant bias in eyewitness identification lineups. *Law Hum. Behav.* 5:299–309

Malpass RS, Devine PG. 1981. Eyewitness identification: lineup instructions and the absence of the offender. *J. Appl. Psychol.* 66: 482–89

McConkey KM, Roche SM. 1989. Knowledge of eyewitness memory. *Aust. Psychol.* 24: 377–84

Meissner C, Brigham JC. 2001. Twenty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychol. Public Policy Law* 7:3–35

Navon D. 1992. Selection of foils by similarity to suspect is likely to misfire. *Law Hum. Behav.* 16:575–93

Noon E, Hollin CR. 1987. Lay knowledge of eyewitness behaviour: a British survey. *Appl. Cogn. Psychol.* 1:143–53

Phillips MR, McAuliff BD, Kovera MB, Cutler BL. 1999. Double-blind photoarray administration as a safeguard against investigator bias. *J. Appl. Psychol.* 84:940–51

Powers PA, Andriks JL, Loftus EF. 1979. Eyewitness accounts of males and females. *J. Appl. Psychol.* 64:339–47

Pozzulo JD, Lindsay RCL. 1998. Identification accuracy of children versus adults: a meta-analysis. *Law Hum. Behav.* 22:549–70

Pozzulo JD, Lindsay RCL. 1999. Elimination lineups: an improved identification procedure for child eyewitnesses. *J. Appl. Psychol.* 84:167–76

Read JD, Vokey JR, Hammersley R. 1990. Changing photos of faces: effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *J. Exp. Psychol.: Learn. Mem. Cogn.* 16:870–82

Risinger MD, Saks MJ, Thompson WC, Rosenthal R. 2002. The Daubert/Kumho implications of observer effects in forensic science: hidden problems of expectation and suggestion. *Calif. Law Rev.* 90:1–56

Robinson MD, Johnson JT, Herndon F. 1997. Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *J. Appl. Psychol.* 82:416–25

Rosenthal R. 1976. *Experimenter Effects in Behavioral Research.* New York: Irvington

Ross DF, Read JD, Toglia MP, eds. 1994. *Adult Eyewitness Testimony: Current Trends and Developments.* New York: Cambridge Univ. Press

Scheck B, Neufeld P, Dwyer J. 2000. *Actual Innocence.* New York: Random House

Seelau SM, Wells GL. 1995. Applied eyewitness research: the other mission. *Law Hum. Behav.* 19:317–22

Shapiro PN, Penrod SD. 1986. Meta-analysis of racial identification studies. *Psychol. Bull.* 100:139–56

Shaw JI, Skolnick P. 1999. Weapon focus and gender differences in eyewitness accuracy: arousal versus salience. *J. Appl. Soc. Psychol.* 29:2328–41

Shaw JS III. 1996. Increases in eyewitness confidence resulting from postevent questioning. *J. Exp. Psychol.: Appl.* 2:126–46

Shaw JS III, McClure KA. 1996. Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law Hum. Behav.* 20:629–54

Smith SM, Lindsay RCL, Pryke S. 2000. Postdictors of eyewitness errors: Can false identification be diagnosed? *J. Appl. Psychol.* 85:542–50

Sporer SL. 1992. Post-dicting eyewitness accuracy: confidence, decision times and person descriptions of choosers and non-choosers. *Eur. J. Soc. Psychol.* 22:157–80

Sporer SL. 1993. Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *J. Appl. Psychol.* 78:22–33

Sporer SL. 1994. Decision times and eyewitness identification accuracy in simultaneous and sequential lineups. See Ross et al. 1994, pp. 300–27

Sporer SL, Penrod SD, Read JD, Cutler BL. 1995. Choosing, confidence, and accuracy: a meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychol. Bull.* 118:315–27

Steblay NM. 1992. A meta-analytic review of the weapon focus effect. *Law Hum. Behav.* 16:413–24

Steblay NM. 1997. Social influence in eyewitness recall: a meta-analytic review of lineup instruction effects. *Law Hum. Behav.* 21:283–98

Steblay NM, Dysart J, Fulero S, Lindsay RCL. 2001. Eyewitness accuracy rates in sequential and simultaneous lineup presentations: a meta-analytic comparison. *Law Hum. Behav.* 25:459–74

Tech. Work. Group Eyewitness Evidence. 1999. *Eyewitness Evidence: A Guide for Law Enforcement.* Washington, DC: US Dep. Justice, Off. Justice Programs

Tollestrup PA, Turtle JW, Yuille JC. 1994. Actual witnesses to robbery and fraud: an archival analysis. See Ross et al. 1994, pp. 144–62

Wells GL. 1978. Applied eyewitness testimony research: system variables and estimator variables. *J. Pers. Soc. Psychol.* 36:1546–57

Wells GL. 1984a. The psychology of lineup identifications. *J. Appl. Soc. Psychol.* 14:89–103

Wells GL. 1984b. How adequate is human intuition for judging eyewitness testimony? In *Eyewitness Testimony: Psychological Perspectives*, ed. GL Wells, EF Loftus, pp. 256–72. New York: Cambridge Univ. Press

Wells GL. 1993. What do we know about eyewitness identification? *Am. Psychol.* 48:553–71

Wells GL. 2001. Police lineups: data, theory, and policy. *Psychol. Public Policy Law* 7:791–801

Wells GL, Bradfield AL. 1998. "Good, you identified the suspect:" Feedback to eyewitnesses distorts their reports of the witnessing experience. *J. Appl. Psychol.* 83:360–76

Wells GL, Bradfield AL. 1999. Distortions in eyewitnesses' recollections: Can the post-identification feedback effect be moderated? *Psychol. Sci.* 10:138–44

Wells GL, Ferguson TJ, Lindsay RCL. 1981. The tractability of eyewitness confidence and its implication for triers of fact. *J. Appl. Psychol.* 66:688–96

Wells GL, Hryciw B. 1984. Memory for faces: encoding and retrieval operations. *Mem. Cogn.* 12:338–44

Wells GL, Leippe MR. 1981. How do triers of fact infer the accuracy of eyewitness identifications? Memory for peripheral detail can be misleading. *J. Appl. Psychol.* 66:682–87

Wells GL, Leippe MR, Ostrom TM. 1979. Guidelines for empirically assessing the fairness of a lineup. *Law Hum. Behav.* 3:285–93

Wells GL, Lindsay RCL. 1980. On estimating the diagnosticity of eyewitness nonidentifications. *Psychol. Bull.* 88:776–84

Wells GL, Lindsay RCL, Ferguson TJ. 1979. Accuracy, confidence, and juror perceptions in eyewitness identification. *J. Appl. Psychol.* 64:440–48

Wells GL, Luus CAE. 1990. Police lineups as experiments: Social methodology as a framework for properly conducted lineups. *Personal. Soc. Psychol. Bull.* 16:106–17

Wells GL, Malpass RS, Lindsay RCL, Fisher RP, Turtle JW, Fulero S. 2000. From the lab to the police station: A successful application of eyewitness research. *Am. Psychol.* 55:581–98

Wells GL, Olson EA. 2002. Eyewitness identification: Information gain from incriminating and exonerating behaviors. *J. Exp. Psychol.: Appl.* In press

Wells GL, Rydell SM, Seelau EP. 1993. On the selection of distractors for eyewitness lineups. *J. Appl. Psychol.* 78:835–44

Wells GL, Seelau EP, Rydell SM, Luus CAE. 1994. Recommendations for properly conducted lineup identification tasks. See Ross et al. 1994, pp. 223–44

Wells GL, Small M, Penrod SD, Malpass RS, Fulero SM, Brimacombe CAE. 1998. Eyewitness identification procedures: recommendations for lineups and photospreads. *Law Hum. Behav.* 22:603–7

Wells GL, Turtle JW. 1986. Eyewitness identification: the importance of lineup models. *Psychol. Bull.* 99:320–29

Wogalter MS, Marwitz DB, Leonard DC. 1992. Suggestiveness in photospread lineups: Similarity induces distinctiveness. *Appl. Cogn. Psychol.* 6:443–53

Wright DB, McDaid AT. 1996. Comparing system and estimator variables using data from real lineups. *Appl. Cogn. Psychol.* 10:75–84