What Can Psychology Say About the Neil v. Biggers Criteria for Judging Eyewitness Accuracy?

Gary L. Wells and Donna M. Murray
University of Alberta, Edmonton, Alberta, Canada

In an influential case (Neil v. Biggers, 1972), the U.S. Supreme Court identified five criteria to be considered in judging eyewitness identification evidence. The Court's intuitions may need reassessment in light of recent psychological research evidence. Especially in need of review is the assumption that eyewitness certainty and eyewitness identification accuracy are usefully related. Under a variety of procedures that mimic the real-world environment, experiments have found the relationship between eyewitness accuracy and confidence to be of little or no practical use. In addition, the Court's assumption that the degree of the eyewitness's attention is critical to judging the accuracy of the eyewitness can be refined for the trier of fact. Psychological research can also bear directly on refining the Court's reference to the witness's opportunity to view the criminal, especially with regard to systematic biases in witnesses' estimations of temporal duration. The Court's reference to an eyewitness's accuracy of prior description as a cue to accuracy of identification is questionable. Finally, the Court's reference to decay of memory across time should be refined to include concepts of interference (i.e., consideration of retention interval events rather than time, per se). Although the Court's intuitions about eyewitness identification are not unreasonable, the recent growth of a literature using forensically relevant procedures has increased the gap between experimental evidence and the Biggers criteria.

"Identifying the defendant as the wrong doer presents an issue, and often the sole one for determination, in every criminal trial" (Wooscher, 1977, p. 969). Among the most persuasive ways to establish the defendant's guilt is eyewitness identification evidence. Eyewitness identification is direct evidence, rather than circumstantial, and undoubtedly has accounted for innumerable convictions of guilty persons who would not have been successfully prosecuted without such evidence. Conversely, many authors have compiled case histories of miscarriages of justice resulting from mistaken identification (e.g., Block, 1963; Borchard, 1932; Frankfurter, 1927; Gardner, 1952), and recently experimental psychologists have taken systematic approaches to delineating problems in the psychology of eyewitness testimony (e.g., see Loftus, 1979). The courts are not oblivious to the issue. The U.S. Supreme Court, for example, has noted that "the vagaries of eyewitness identification are well-known, the annals of criminal law are rife with instances of mistaken identification" (United States v. Wade).

Although the courts are aware of the general problem of mistaken identification, there is some confusion over how to deal with the issue. Recent years have seen an increased presence of psychologists in court who provide expert testimony regarding eyewitness evidence. But the courts are inconsistent in allowing such testimony, and some psychologists have argued that allowing expert testimony at individual trials might not be the level at which the issues should be handled (e.g., Egeth & McCloskey, 1983; Wells, 1978). Furthermore, expert testimony in individual trials is a costly and time-consuming process. Something on a larger scale might prove more efficient, such as instructions to juries deliv-

The authors thank John Brigham, Kenneth Deffenbacher, Elizabeth Loftus and Leendert Mos for helpful comments on a manuscript version of this article.

Requests for reprints should be sent to Gary L. Wells, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9.

1 United States v. Wade, 388 U.S. 218 (1967)
ered routinely by judges. Expert testimony might still be used, of course, in particularly difficult cases.

This article begins with a background of the current legal position on eyewitness evidence, followed by a discussion of each of the criteria, which the Court has identified, for evaluating eyewitness accuracy. Currently, there are five criteria set forth by the U.S. Supreme Court that seem to guide most lower courts.

**Legal Background**

The official position of the U.S. Judiciary with respect to eyewitness identification is derived from a 1972 U.S. Supreme Court decision (*Neil v. Biggers*). In this case, the Court discussed the factors to be considered in determining whether the pretrial eyewitness identification has violated due process. The due process rule as a sanction is applicable to many areas of criminal law and is derived, of course, from the Fourteenth Amendment. Prior to *Biggers*, the general standard for the due process rule in eyewitness identifications concerned the suggestiveness of the identification procedure (*Stovall v. Denno*). The *Neil v. Biggers* case, however, shifted the emphasis from suggestiveness to accuracy. Previously, overly suggestive identification procedures meant that the court must suppress evidence of the pretrial identification. For example, a "show-up" (one-to-one confrontation between an eyewitness and a suspect) is generally considered overly suggestive; and, under the old guidelines, if there was reasonable opportunity to use a lineup, a show-up violated due process. In the *Neil v. Biggers* case (wherein a show-up was at issue), however, the Court argued that "it is the likelihood of misidentification which violates the defendant's right to due process" and that the crucial issue is whether "the identification procedure was reliable even though the confrontation procedure was suggestive." (p 198-199) (Note: The Court uses the word 'reliability' to refer to what researchers would call validity or accuracy in this context.)

The Court listed five factors in the *Neil v. Biggers* case to be considered in determining accuracy:

1. The opportunity of the witness to view the criminal, at the time of the crime.
2. The witness' degree of attention.
3. The accuracy of the witness' prior description of the criminal.
4. The level of certainty demonstrated by the witness at the time of confrontation.
5. The length of time between the crime and the confrontation.

Thus, the Court went beyond the issue of due process as it had been treated previously and specified that the true issue is one of accuracy. Some confusion has resulted from this as some legal experts have been unable to see how these five factors constitute due process considerations (e.g., *State v. McManus*; Sobel & Pridgen, 1981). It could be said that the Court has offered an intuitive theory of eyewitness identification or at least a set of five hypotheses.

Since 1972, lower courts appear to have adopted the *Biggers* standards of eyewitness accuracy to the point of even rejecting due process concerns (e.g., the suggestibility of the identification procedure, see, e.g., *State v. Henderson*; *State v. Bono*, *United States v. Anderson*, *Haberstrok v. Montanye*, *Kirby v. Sturges*, *Israel v. Odom*, and *United States v. Rodriguez*). These courts have operated as though such factors as certainty, opportunity to view the criminal at the time of the crime, accuracy of prior description, attention by the witness, and/or the temporal duration between crime and confrontation are compelling criteria that can overcome traditional concerns about due process (e.g., suggestiveness of the identification procedure). Indeed, this is what the Court meant when it rejected a per se exclusionary rule in the case of *Manson v. Braithwaite*. This case...

---

3 *Stovall v. Denno*, 388 U.S. 293 (1967)
4 *State v. McManus*, 263 N.W.2d 556, 559 (Iowa, 1977)
7 *United States v. Anderson*, 490 F.2d 785 (D.C. Cir. 1974)
8 *Haberstrok v. Montanye*, 493 F.2d 483 (2d Cir. 1974)
9 *Kirby v. Sturges*, 510 F.2d 397 (7th Cir. 1975)
10 *Israel v. Odom*, 521 F.2d 1370 (7th Cir. 1975)
11 *United States v. Rodriguez*, 510 F.2d 1 (5th Cir. 1975)
revolved around a narcotics agent's identification of Braithwaite from a single photograph (i.e., there were no alternative photographs), a procedure which the prosecution admitted was both suggestive and unnecessary. The Supreme Court approved the use of the Biggers's criteria, in doing so, however, the Court made the suggestive identification procedure irrelevant. The Court pointed to the fact that the agent was exposed to the person in question for 2 to 3 minutes at the time of the drug deal, the agent was a trained observer who was therefore likely to attend to detail, the agent's verbal description was detailed, the agent was certain about the identification, and only 2 days elapsed between the incident and the identification.

The impact of the Biggers case raises a number of questions about the adequacy of the five criteria. The Court has relied on intuition to specify some factors that it believes are correlated with eyewitness identification accuracy. Except for a dissenting opinion by Justices Brennan and Marshall in Watkins v. Sowders, the Court has not made use of experimental psychology in dealing with matters of eyewitness testimony. It could be argued that it was unwise of the Court to fail to make use of research in experimental psychology. After all, psychologists have been studying memory in a scientific way for nearly 100 years (e.g., Ebbinghaus, 1885/1964). This is especially true today, as there is an empirical literature specifically on the psychology of eyewitness testimony that is large enough to justify five recent book-length volumes (Clifford & Bull, 1978, Davies, Ellis, & Shepherd, 1981, Loftus, 1979, Wells & Loftus, 1983, Yarmey, 1979), a special issue of Law and Human Behavior (Vol. 4, 1980) devoted to empirical studies of eyewitness testimony, and a recent conference devoted entirely to the psychology of eyewitness testimony (Alberta Conference on the Psychology of Eyewitness Testimony, 1980). Thus, advice to the Court regarding eyewitness matters would seem to be much more viable on the basis of empirical evidence today than it was in 1972.

Each of the factors identified by the Court in Neil v. Biggers (hereafter called the Biggers criteria) is addressed with regard to modern evidence in the experimental psychology of eyewitness testimony. It is not the case that each of these five factors has been equally researched and none have been plenarily researched. Indeed, one of the purposes of this article is to point out weaknesses in current psychological evidence on these issues so that psychology will be prepared to provide useful information to the Court should they ask for such in a future case. A second purpose is to help draw attention to any discrepancies between legal opinion and current empirical evidence in psychology.

### Eyewitness Certainty

#### Psychological Evidence

Eyewitness identification certainty, defined as the degree to which an eyewitness believes that his or her identification is accurate, has received considerable experimental attention in recent years. Prior to the mid-1970s there was little in the published psychological literature that would lead one to suspect that eyewitness certainty would be a poor or misleading cue to eyewitness accuracy. Why the question was not put to empirical test prior to that time is not clear. Perhaps it stemmed from an assumption that accuracy-confidence relationships in other tasks and settings (e.g., Brown & McNeill, 1966; Kozlowski & Bryant, 1977) would generalize to eyewitness identifications. Perhaps the intuition (that accuracy and confidence are related in eyewitness identifications) is so compelling that empirical tests had no a priori interest value. Whatever the explanation, relevant experiments have now been conducted and they do not support the intuition that eyewitness certainty should be considered a good index to eyewitness accuracy.

The experiments bearing on this issue have involved a number of different laboratories, subject populations, and procedures. Some studies required identification from facial photographs (e.g., Brown, Deffenbacher, & Sturgill, 1977), whereas others required identification from live lineups (e.g., Malpass & Devine, 1981a, 1981b). In some studies the original encounter with the perpetrator occurred via videotaped or slide scenarios (e.g.,

---

Clifford & Scott, 1978; Loftus, Miller, & Burns, 1978), whereas in other studies the original encounter occurred during a staged crime (e.g., Leippe, Wells, & Ostrom, 1978; Sanders & Warnick, 1981). Some studies have led the witness to believe the confrontation is real and that identification has serious consequences for the accused (e.g., Malpass & Devine, 1980, Sanders & Warnick, 1981; Murray & Wells, 1982), whereas others have informed the eyewitness (after the original incident) that the identification is part of an experiment (e.g., Buckhout, Alper, Chern, Silverberg, & Slomovits, 1974; Buckhout, Figueroa, & Hoff, 1975; Lindsay & Wells, 1980, Wells & Leippe, 1981). Some of the studies measured eyewitness certainty at the time of identification with Likert-type scales (e.g., Clifford & Scott, 1978; Davies, Ellis, & Shepard, 1978; Leippe et al., 1978), whereas other studies have measured open-ended statements of certainty (e.g., Murray & Wells, 1982) Finally, some studies have measured certainty-related phenomenon, such as willingness to undergo cross examination (e.g., Lindsay, Wells, & Rumpel, 1981, Wells, Lindsay, & Ferguson, 1979) and persuasive impact of testimony (e.g., Wells, Ferguson, & Lindsay, 1981), to see their relationship to eyewitness accuracy With rare exceptions, these studies show that the certainty of those eyewitnesses who make an identification accounts for no more than 4% or 5% of the variation in their accuracy; most of the studies show the relationship to be statistically nonsignificant Under some conditions, such as perpetrator-absent lineups or misleading postevent information, the accuracy–confidence correlation is actually negative (e.g., Malpass & Devine, 1981a; Loftus et al., 1978). The strength and direction of the confidence–accuracy relationship has been shown to be affected by three variables: the presence/absence of the perpetrator in the lineup, the presence/absence of variables that affect only one member of the confidence–accuracy pair, and the consequences of the identification Each of these is discussed in turn.

Malpass and Devine (1981a) exposed unsuspecting eyewitnesses to a staged vandalism followed by a live lineup. By varying the lineup so that it included or did not include the vandal, Malpass and Devine were able to show that eyewitness certainty was related primarily to the act of making an identification choice, rather than the accuracy of the choice. When the perpetrator was absent from the lineup, for example, those who made a choice were more certain that they were correct than were those who did not make a choice (even though all of the former were inaccurate and all of the latter were accurate) The negative correlation between certainty and accuracy in the perpetrator-absent conditions was as robust as was the positive correlation in the perpetrator-present conditions, as a result the overall correlation was functionally zero. Because we cannot assume in the real world that a given lineup contains the true criminal, the Malpass and Devine study suggests that the certainty measure may be useless as a predictor of accuracy and calls into question those confidence–accuracy correlations that were obtained from studies that used only perpetrator-present lineups (e.g., Brigham, Mass, Snyder, & Spaulding, 1982, Malpass & Devine, 1981b).

There are theoretical reasons to believe that various events can affect identification accuracy without affecting identification confidence and vice versa (see Leippe, 1980). These theoretical expectations are supported by several eyewitness experiments that show that accuracy can be altered without any corresponding modification of confidence (e.g., Lindsay & Wells, 1980, Lindsay et al., 1981, Loftus et al., 1978) The Lindsay and Wells study varied the physical similarity of lineup members to each other and found strong effects on the identification accuracy of eyewitnesses to a staged crime. There were no effects, however, on the eyewitnesses' certainty (and no accuracy–certainty correlation in any of the conditions) Lindsay et al varied the conditions under which a staged theft was witnessed and found strong effects on the identification accuracy of eyewitnesses Conditions of poorest accuracy, however, did not differ from conditions of best accuracy in terms of eyewitness certainty. Loftus et al. varied the presence/absence of misleading postevent information and found that the presence of misleading information greatly reduced memory accuracy but increased, rather than decreased, certainty on the part of the eyewitnesses that
their memories were accurate. Thus, some things affect accuracy without corresponding effects on certainty. As a result, accuracy and certainty lose calibration with each other.

In addition to evidence that accuracy can be affected without corresponding changes in certainty, a recent study by Wells et al (1981) showed that certainty can be affected by things that do not affect accuracy. In this study, eyewitnesses to a staged theft made an attempt to identify the thief from a six-person photo lineup and were then either warned or not warned to prepare themselves for testimony and cross-examination. The warned eyewitnesses later exhibited higher levels of certainty that their identification was accurate than did unwarned eyewitnesses. This effect was independent of true accuracy. Thus, eyewitness certainty can be affected by factors that do not affect eyewitness accuracy, which further contributes to the noncorrespondence between certainty and accuracy.

A final study that is particularly informative regarding the accuracy-certainty issue was conducted by Murray and Wells (1982). Unsuspecting eyewitnesses were exposed to a staged theft and later attempted identifications of the thief from a six-person photo array. Half of the eyewitnesses were informed after the incident, but before the identification, that the theft was simulated. The other half of the witnesses were led throughout the identification task to believe that the event was real. Eyewitnesses who had knowledge that the theft was simulated showed a strong confidence-accuracy correlation when a multiple regression technique was used. However, this relationship was eliminated (and made slightly negative) in the conditions where the eyewitnesses thought that the theft was real. Thus, believing that the identification has real consequences for the accused seems to interfere with the confidence-accuracy relationship owing, perhaps, to the extra anxiety created in these realistic conditions.

Interestingly, the number of years experience for judges does not relate to their beliefs about the relationship between eyewitness confidence and accuracy, lineup conductors, however, are more likely to be aware of the noncorrespondence between eyewitness confidence and accuracy the more years experience that they have (Brigham & Wolfskel, Note 1). Presumably this is due to the fact that lineup conductors experience situations over the years in which eyewitnesses identify police detectives (or other known innocent foils) from a lineup with high levels of confidence. Judges, however, have no opportunity to observe such occurrences.

Conclusion

The experimental literature does not support the Court's assumption about the utility of eyewitness certainty. It is true that there are conditions under which accuracy and certainty in eyewitness identifications may be related (such as with perpetrator-present lineups). However, because many factors that influence accuracy do not affect certainty and because certainty is influenced by factors that do not relate to accuracy, eyewitness certainty can be misleading. Under some conditions (such as perpetrator-absent lineups, misleading postevent information, or rehearsal of false memories), high levels of eyewitness certainty have been associated with inaccuracy rather than accuracy (see Malpass & Devme, 1981a, Loftus et al., 1978, Wells et al., 1981, respectively, on these points). Studies that mimic the anxiety associated with confrontations experienced by eyewitnesses seem especially likely to guarantee noncorrespondence between accuracy and certainty (e.g., Murray & Wells, 1982; Sanders & Warnick, 1981) There have been three reviews of the accuracy-certainty issue in eyewitness testimony; all three reached the conclusion that the courts should not rely on certainty to infer accuracy (Deffenbacher, 1980, Leippe, 1980, Wells & Murray, 1983). Although Deffenbacher argued that there may be certain optimal conditions in which confidence-accuracy relationships may hold, what constitutes an "optimal condition" is unclear. More research is needed to see if such conditions exist and, if so, whether these conditions apply to real-world eyewitnessing. In addition, research on confidence-accuracy relationships in eyewitnessing has primarily focused on between-subjects comparisons (e.g., Is a confident eyewitness more accurate than a nonconfident eyewitness?) rather than within-subjects comparisons (e.g.,
Is a witness more likely to be accurate when confident than that same witness is when not confident?). Although between-subjects comparisons may be more relevant to the forensic issue as it typically is encountered, within-subjects analyses may also have forensic relevance and should be assessed in further research.

Degree of Attention

It is not clear what the Court means precisely by the term "degree of attention." If by it the Court means how long the witness looked at the perpetrator, then it would seem to be too narrowly defined. Clearly, exposure time is one of the most important factors influencing memory. However, during the last decade there have been impressive demonstrations in the experimental memory literature showing that the amount of time spent attending to visual information can be much less critical than the type of attention. A modern framework for understanding memory processes in this regard is Craik and Lockhart's (1972) levels of processing conceptualization. Studies consistently find that the accuracy of memory retrieval depends on what Craik and Lockhart have called the "depth" to which the original event was processed. Depth is determined by a number of things such as the purpose for attending to the information. Thus, a person who examines words in order to make a superficial judgment (e.g., number of syllables) will not find these words to be as accessible in memory as would a person who judged the words for their meaning (e.g., Aruckle & Katz, 1976; Craik & Tulving, 1975).

The type of attention, rather than time per se, plays a critical role in memory for faces as well (e.g., Bower & Karlin, 1974; Mueller, Bails & Goldstein, 1979, Strnad & Mueller, 1977). Attention to specific facial features, for example, requires a considerable amount of time but is not as effective for memory as is a task that requires deeper processing, such as making a personality trait judgment (e.g., Patterson & Baddeley, 1977; Warrington & Ackroyd, 1975, Winograd, 1976). How the trier of fact determines the type of attention paid by the eyewitness is an open question. However, as discussed in the next section, the same problem exists for assessing the amount of time that the eyewitness attended to the perpetrator. In both cases it seems that one must rely on the eyewitness's report. This could be problematic because the witness easily could err in such reports.

Recent research indicates that it might be important to clarify another aspect of attention for the trier of fact. Specifically, humans have a limited capacity for processing information. Therefore, attention to one stimulus will result in a reduction of the available processing capacity for other stimuli (Kahneman, 1973). Although "holistic processing" can occur under certain circumstances (e.g., see Navon & Gopher, 1979), negative correlations can exist between performance on central versus peripheral tasks (e.g., Hagen, Meacham & Mesibov, 1970). Regarding the latter, recent research using staged thefts shows that the accuracy of eyewitnesses' memories for trivial, peripheral details (e.g., whether there was a window in the room where the theft occurred) can be negatively correlated with the eyewitnesses' accuracy in identifying the thief from a photo lineup (Wells & Leippe, 1981). Subject/jurors in the Wells and Leippe study, however, behaved as though the correlation was positive. That is, they were more likely to believe that the eyewitness had correctly identified the thief if the eyewitness had accurately recalled the peripheral details surrounding the event. Thus, triers of fact seem to approach the eyewitness issue as though perceptual processing were holistic in settings where eyewitnesses are not using holistic processing. The general notion of degree of attention does nothing to caution triers of fact against a supposition of holistic processing. Yet, as Wells and Leippe demonstrated, such a supposition, in conjunction with attorneys' normal attempts to discredit eyewitnesses for failing to notice such peripheral details, produces considerable errors on the part of triers of fact.

Conclusion

Although the Court is not mistaken in suggesting that the degree of attention is a factor...
in accuracy of memory, theoretical and empirical evidence points to the type of attention or level of processing as being more crucial for memory. Unfortunately, research has not yet been conducted that directly addresses this issue in a forensically valid manner. The courts need to know whether an eyewitness's self-report of degree of attention or of type of attention is more diagnostic of accuracy. After all, the levels of processing research manipulates the type of processing to find its effects whereas the courts can only measure self-reports of this factor. We can not assume a priori that self-reports of type of processing will have predictive utility. The empirical test seems to rest on whether a question such as "What were you thinking about while viewing the perpetrator's face?" is more diagnostic of accuracy than is a question such as "Did you attend to the perpetrator's face?"

Attorneys who cross-examine eyewitnesses generally try to discredit the eyewitness by showing that he or she does not recall accurately some trivial, peripheral detail of the event in question. Research shows that such discrediting attempts are successful but promote errors on the part of the triers of fact. The empirical test seems to rest on whether a question such as "What were you thinking about while viewing the perpetrator's face?" is more diagnostic than a question such as "Did you attend to the perpetrator's face?"

Opportunity to View the Assailant

An eyewitness's opportunity to view the assailant at the original encounter is obviously important for memory. There is certainly nothing in experimental psychology that would contradict this global statement. There is good experimental evidence, however, that could refine this general statement

It is particularly important to note that the trier of fact's understanding of the eyewitness's opportunity to view the assailant often depends on the eyewitness description of his or her viewing opportunity. For example, the eyewitness is the one who estimates his or her distance from the assailant, the eyewitness must be the one who describes temporary occlusions of his or her visual contact with the assailant at the time of witnessing. The courts need to know whether an eyewitness's self-report of degree of attention or of type of attention is more diagnostic of accuracy. After all, the levels of processing research manipulates the type of processing to find its effects whereas the courts can only measure self-reports of this factor. We can not assume a priori that self-reports of type of processing will have predictive utility. The empirical test seems to rest on whether a question such as "What were you thinking about while viewing the perpetrator's face?" is more diagnostic of accuracy than a question such as "Did you attend to the perpetrator's face?"

Facial recognition research does not conflict with Justice Blackmun's hypothesis. Laughery, Alexander, and Lane (1971), for example, presented subjects with four slides of faces for a total of 10 sec or 32 sec. These slides were then mixed into a set of 150 slides of faces, and subjects correctly identified the original faces 47% of the time with the 10-sec exposure and 58% of the time with the 32-sec exposure. Thus, short exposures can yield reasonable levels of memory for faces.

14 People v Miller, 43 A D 2d 605, 348 N YS 2d 584 (1973)
16 People v Hahn, 39 Ill App 3d 969 (1976)
17 United States v Lustig, 555 F2d 737 (1977)
18 People v Griswold, 54 Ill App 3d 246, 369 N E 2d 392 (1977)
19 Gov of Canal Zone v Waldron, 574, F2d 283 (1978)
20 Tate v State, 265 S E 2d 818 (1980)
21 Moore v Illinois, 434 U S 220 (1977)
if the circumstances are good. Longer exposures, of course, are usually superior to shorter exposures. What Justice Blackmun and other people may fail to appreciate, however, is the fact that the temporal length of events is generally overestimated by witnesses (e.g., Buckhout et al., 1974, Marshall, 1966; Shifman & Bobko, 1974). This finding is one of the earliest reliable findings to be documented in the psychology of eyewitness testimony (Münsterberg, 1908). The overestimation tendency is even more pronounced when the witness is feeling stress or anxiety (Sarason & Stroops, 1978). One should, therefore, question whether a witness's 10-sec estimate is accurate because in general the only way estimates of time can be obtained is by asking the witness.

Thus, although eyewitness research cannot pin down precisely an amount of exposure time that is sufficient for good memory (it depends on too many other factors), there is solid evidence that exposure time estimates from eyewitnesses should be scaled down. Exceptions would apply, of course, in cases where the witness had external validators of temporal duration to which he or she was attending (e.g., a clock).

A study that we recently conducted indicates that witnesses also tend to underestimate the proportion of time that their visual contact with a person is occluded by another object. In this experiment, 80 persons individually viewed a videotaped scenario lasting 5 minutes. They were asked to attend to a specific person in a crowded room whose image was fully occluded for 20%, 50%, or 80% of the total 5 minutes. The observers were later asked to estimate the percentage total time that the target person was not in clear view. Although the observers were sensitive to occlusion time, their mean estimates in the three conditions were 10.6%, 21.3%, and 53.5%, respectively. The estimations of occlusion time were statistically lower ($p < .05$) than were actual occlusion times in all three conditions. Also, the average estimate for length of the videotape scenario was 9.2 minutes (cf. the actual time of 5 minutes). Thus, in addition to overestimating total time for the scenario, subjects underestimated occlusion times, as a result, the 80% occlusion condition (1 minute of actual exposure to target) was estimated to be almost 4½ minutes of exposure to the target.

**Conclusion**

The critical point of this issue is that information regarding a witness's opportunity to view the assailant is generally obtainable only via reports of the very witness whose reliability is yet to be assessed. This state of affairs is not hopelessly circular, however, because there are some aspects of opportunity for viewing reports (e.g., temporal durations) that are associated with systematic biases. The types of errors that eyewitnesses make in describing their witnessing conditions needs considerably more research. Representative research on this matter is as yet too narrow to be of much use in helping the Court refine its judgments about the notion of opportunity to view the assailant. In particular, more research will have to be conducted on the question of how well people's self-reports of witnessing conditions predicts their real witnessing conditions, since self-report is the only information available to triers of fact.

**Accuracy of Prior Description**

In *Neil v Biggers* the Court listed "accuracy of the witness' prior description of the criminal" as one of the five factors that indicate the accuracy of subsequent identification. In practice, of course, one cannot assess the accuracy of the witness's prior description of the criminal unless one assumes that the defendant is guilty. What the Court really meant to focus on was the degree of similarity between the witness's prior description of the criminal and the defendant's physical characteristics. If the defendant is guilty, of course, the two statements (i.e., accuracy in prior description of criminal and similarity between prior description and suspect characteristics) are identical. However, the guilt or innocence of the defendant is the question under consideration and cannot be presumed. Thus, the prior description criterion should be rephrased so that it refers to similarity between prior description and defendant characteristics.

It might not be adequate to simply clear
up this logical anomaly regarding the prior description criterion. Specifically, it is not clear how much diagnostic value there is in the match between an eyewitness's prior verbal description of the criminal and the eyewitness's choice of a lineup member. First, a number of studies have shown that people who are superior at naming details of faces or verbally describing them from memory are not appreciably better at recognizing faces (e.g., Goldstein, Johnson, & Chance, 1979, Howells, 1938; Wolfskiel & Brigham, Note 2). Another factor that must be considered here is the fact that verbal descriptions and recognition of faces can be correlated for spurious reasons (i.e., other than accuracy). A study by Loftus and Greene (1980), for example, led witnesses in some conditions to incorrectly remember an assailant as having a moustache or curly hair. Subsequent testing showed that these witnesses incorporated the moustache or curly hair into their verbal descriptions and that the witnesses chose someone from a photo array who had such features. Thus, errors in the witnesses' prior descriptions also carried over to errors in their identifications. The point is that a prior description and a subsequent recognition can corroborate each other when in fact both the description and recognition are in error. In addition, an individual's prior description to some extent determines the content of the lineup, thereby providing a route through which errors in the original description can reoccur in the lineup identification. Obviously, this becomes increasingly important as the lineup becomes increasingly suggestive.

The Court's reference to the eyewitness's accuracy of prior description as a clue to the eyewitness's accuracy of identification also relates to a long-standing issue in memory research, namely recognition versus recall. The fact is that recall and recognition often yield different results. Usually recognition seems to be superior to recall. Apparently this is due to the fact that the presence of the target item facilitates access to the stored information via activating retrieval cues (Brown, 1976). On the other hand, there are situations in which recognition can fail in spite of an individual's ability to recall (Tulving & Watkins, 1977). The latter phenomenon is likely to occur whenever the retrieval information (context) provided at recall is better than that provided by the recognition test. Perhaps most relevant to the current concern is recent evidence that retrieval information used in recall is uncorrelated with that used in recognition (Broadbent & Broadbent, 1977; Flexser & Tulving, 1978). It appears that in recall some aspect of the context is provided and the individual must retrieve the target information, whereas in recognition the target information is provided and the individual must retrieve contextual aspects of the original episode.

The finding that recognition probability for a word is unaffected by whether or not a word is recallable (Flexser & Tulving, 1978) portends poorly for using recall to predict recognition. Unfortunately, this research has not been conducted with faces, owing in part to the difficulty in operationally defining accuracy of face recall. Nevertheless, current data and theory do not lend much credibility to the Court's emphasis on accuracy of prior description as a criterion for judging accuracy of criminal identification by eyewitnesses.

Conclusion

It is true, unfortunately, that the issue of similarity between prior description and subsequent recognition is, and will remain, a difficult question to address empirically. The main reason for this is that a suspect is placed in a lineup, at least in part, because he or she matches the prior description. This creates a curious situation of unknown and varying contingencies for generating suspects that may be difficult to simulate in controlled research settings. In addition, the current literature on recall and recognition memory suggests that independent processes are involved. Finally, three experiments have failed to find any appreciable relationship between a person's prior description of a face and the person's accuracy in identifying the face (Goldstein et al., 1979; Howells, 1938; Wolfskiel & Brigham, Note 2). Nevertheless, further research specifically directed at factors that might produce correspondence/noncorrespondence between eyewitnesses' prior de-
scriptions and subsequent identifications could help clarify the issue.

Time Between Witnessed Event and Test

It is a well-established fact that forgetting increases with time. This general principle is one of the oldest in experimental psychology (Ebbinghaus, 1885/1964), and it made good intuitive sense for the Court to mention it in the Biggers case as one of the criteria for eyewitness accuracy. Yet, the mentioning of time per se can be misleading to those who must use it as a criterion. There are two reasons why the Court may wish to refine its statement about time. First, there may be a strong tendency for people to use their own memory for faces of familiar others as an appropriate comparison. In this regard, memory for faces is extremely stable over time for familiar faces. Bahnck, Bahnck, and Wittlinger (1975), for example, found that people could identify up to 90% of the faces of their high school classmates as many as 35 years after graduation. If a trier of fact thinks of his or her memory for familiar faces when judging the effects of time, he or she is likely to believe that the decay over time of an eyewitness’s memory for the criminal’s face is trivial.

A setting that is more forensically relevant involves the rate at which memory decays for the face of a person seen only once before. A recent study by Shepherd, Davies, and Ellis (Note 3) bears on this issue. In this study unsuspecting eyewitnesses observed a staged incident enacted by one of four target persons. Identification parades using nine persons (one of which was the target person) were administered at delays of 1 week, 1 month, 3 months, and 12 months. Accuracy of identification was highly stable over the 1-month period (about 10%) but dropped to a level expected by chance by the 12-month period.

Although most studies testing recognition memory for faces seen only once before appear to show that the level of recognition accuracy declines very little over periods up to 30 days (e.g., Chance, Goldstein, & McBride, 1975; Davies et al., 1978; Egan, Pittner, & Goldstein, 1977; Laughery, Fessler, Lenorovitz, & Yoblick, 1974; Shepherd & Ellis, 1973), these studies may be testing the nearly asymptotic portion of the forgetting curve. Indeed, a study by Shepherd and Ellis (1973) found that under optimal laboratory conditions 11% of the information was forgotten just 3 minutes after exposure.

It is well known in psychology that time per se may not be nearly so critical to memory retention as is what occurs during the retention interval. A common factor that can occur during the witness-test interval and that is known to affect the accuracy of identification is the viewing of mugshots (e.g., Davies, Shepherd, & Ellis, 1979; Gorenstein & Ellsworth, 1980; Laughery et al., 1971). In addition, a great deal of eyewitness research has been conducted showing that information acquired between a witnessed event and a subsequent memory test can be incorporated into the witness’s memory report (see Loftus, 1979, Chapter 4 for a review). A recent study by Loftus and Greene (1980) illustrates this point. In this study subjects viewed a person in a photograph (Experiment 2), a film (Experiment 1), or live (Experiment 3) and subsequently encountered a description ostensibly given by another witness. Whenever this other witness referred incorrectly to a feature (e.g., mentioning a moustache when in fact the person had no moustache), nearly 70% of the subjects subsequently identified a person from a 12-person lineup who had that feature (cf 13% in the control group). Research indicates that there are numerous events that can occur between a witnessed event and subsequent memory testing that can have profound effects on memory reports (e.g., see Loftus, 1975, 1977, 1981, Loftus & Palmer, 1974, Loftus & Zanni, 1975).

Conclusion

There is nothing incorrect, technically, about the Court’s listing of the witness–test interval as a factor in accuracy. However, some further guidance for the trier of fact may be in order. For example, the trier of fact should not judge the effects of time by intuiting what happens with familiar faces, which show robust retention over long periods of time. Evidence regarding memory for
faces seen only once before shows little decline up to 30 days but almost total loss by 12 months. Also, the Court should consider events that occur during the witness-test interval rather than emphasizing time per se. There is now considerable evidence that other people's (e.g., the interrogator or other witnesses) verbally conveyed assumptions during the witness-test interval can affect the eyewitness's memory and that mugshot viewing decreases the accuracy of subsequent identifications.

Other Factors

Some courts have refused to consider reliability factors not mentioned in Neil v. Biggers (e.g., State v McManus, see Footnote 4). Other courts have been willing to add other factors to the list. In one case (Davis v Jennings), the fact that some witnesses were not able to identify the defendant in a lineup was considered to be evidence that the lineup was not overly biased. In the Davis v Jennings case the court used nonidentifications by other witnesses to boost the credibility of an identification by a given witness. In effect, the Davis v Jennings case presents an anomalous argument wherein the nonidentifications served as incriminating evidence. As Wells and Lindsay's (1980) mathematical analysis of lineup diagnosticity shows, however, the mathematics of uncertainty reduction as well as empirical data should lead one to conclude that nonidentifications are diagnostic of the suspect's innocence. In addition, recent studies using staged crimes show that biased lineups do not affect the no-identification rate. Instead the lineup biases affect the distribution of identifications so that the probability of identifying the suspect is increased (Lindsay & Wells, 1980). In general, the evidence does not support a reliance on multiple-witness patterns for inferring fairness of the lineup.

In other cases in recent years the courts have considered such factors as the eyewitness's intelligence (e.g., State v St Onge) and the eyewitness's occupation (e.g., Royce v Moore). Studies showing null and even negative relationships between eyewitness intelligence and eyewitness testimony accuracy date back as far as the Aussage researchers as reported by Whipple (1909). As for occupation, police officers are considered by many courts to be more credible as eyewitnesses than are civilians (e.g., United States v Bothwell, United States v Flickinger). Yet, experimental studies typically yield no differences in the accuracy of police versus civilians as eyewitnesses (e.g., Clifford, 1976; Clifford & Richards, 1977; Tickner & Poulton, 1975; Verinis & Walker, 1970), nor do they appear to be any effects for the length of time an individual has been involved in relevant police work (Bilig & Miller, 1976). Although the courts might be justified in considering prior eyewitness training for a given police officer, the effectiveness of such training is yet to be determined. In general, it seems that training and experience make police more likely to perceive criminality or potential criminality in ambiguous scenes but does not make police better at reporting accurately about people and actions (Tickner & Poulton, 1975).

Suggestiveness

The Biggers criteria had the effect of dropping the lower courts' concerns with the suggestiveness of pretrial identification procedures on the argument that the real concern should be one of eyewitness accuracy (see Footnotes 5-11). Although the Court was wise to emphasize accuracy as the critical concern, it does not seem wise to drop suggestiveness as an important factor. Indeed, there is now ample evidence to show that the suggestiveness of identification procedures has profound effects on eyewitness accuracy (Lindsay & Wells, 1980; Loftus & Greene, 1980; Malpass & DeVine, 1981a; Fanselow & Buckhout, Note 4). Prior to the Biggers case, suggestiveness was considered of paramount importance and was

---

22 United States ex rel Davis v Jennings, 414 F Supp 544 (E D Pa 1976)
23 State v St Onge, 392 A2d 47, 51 (Me 1978)
24 Royce v Moore, 469 F2d 808 (1st Cir 1972)
25 United States v Bothwell, 465 F2d 217 (9th Cir 1972)
26 United States v Flickinger, 573 F2d 1349 (9th Cir), cert denied, 439 U S 836 (1978)
most often typified by the use of show-ups (i.e., what is functionally a one-person lineup). A staged-theft study by Wells and Lindsay (1980) shows how this type of suggestiveness affects eyewitness identification accuracy. Lindsay and Wells found that the likelihood of mistaken identification from lineups whose functional size approached a value of 1.0 produced over twice the rate of false identifications as that obtained from lineups whose functional size approached a value of 3.0. Obviously, this effect occurs because the true perpetrator is not always in the lineup and because the lineup detractors are known to be innocent. Thus, both logic and experimental data tell us that the suggestiveness of show-ups or low functional size lineups is not independent of the accuracy issue.

Fanselow and Buckhout (Note 4) operationalized suggestiveness by manipulating nonverbal cues to eyewitnesses as they viewed mugshot arrays. After viewing a filmed mugging, eyewitnesses were shown an array of six facial pictures. The “identification officer” leaned forward and smiled when the eyewitness viewed certain pictures, for other pictures he leaned backward and frowned; and for other pictures he acted neutrally. The leaning cues enhanced false identifications on the pictures with which the cues were associated.

Loftus and Greene (1980) conducted a study in which eyewitnesses were given cues by another person that falsely suggested the presence of a facial characteristic (e.g., mustache) on an actor who had been observed previously. Significant percentages of eyewitnesses who were recipients of such a suggestion incorporated that characteristic into subsequent descriptions of the actor and falsely identified a person from a set of pictures who had such a characteristic.

Finally, suggestiveness has been investigated with regard to the instructions given to eyewitnesses at the time of identification. Malpass and Devine (1981a) staged a vandalism for a large number of unsuspecting eyewitnesses who then viewed a lineup that did or did not contain the vandal. Some witnesses were told that the vandal may or may not be present, whereas it was suggested to others that the vandal was present. The suggestive instructions produced an average error rate of 51.5%, whereas the error rate was only 25.0% with the nonsuggestive instructions.

**Conclusion**

The Biggers criteria have left lower courts with the impression that the suggestiveness of pretrial identification procedures is of little concern. In fact, suggestiveness has robust effects on accuracy, which is what the Biggers’s criteria were designed to address. Suggestiveness is an important factor to consider when judging the accuracy of eyewitness identification evidence and it is importantly distinct from the five criteria outlined in Neil v. Biggers. It is particularly important to note that none of the five Biggers criteria concern themselves with conditions of retrieval. Memory researchers have long known that testing or retrieval conditions have strong effects on the accuracy of memory judgments. If and when the Court reconsiders the Biggers criteria, it should reestablish the importance of retrieval/testing conditions in general and stress the importance of suggestiveness in particular.

**Prospectus**

No general form of advice to triers of fact will allow them to make perfect judgments about the accuracy of a given eyewitness’s account. Recently eyewitness research has involved staging crimes for unsuspecting eyewitnesses and then cross-examining those witnesses (e.g., Wells et al., 1979). It is clear from this research that the task of discerning truth is extremely difficult. Most of the time there is little or no information that can be discerned from testimony that allows for discrimination between accurate and inaccurate testimony; the principal difficulty results from the high certainty of inaccurate eyewitnesses (see Wells et al., 1981). Yet, it is also clear that the courts can improve on the advice/criteria used currently in judging the accuracy of eyewitness identification evidence. If the courts are to have formal criteria at all, a serious review of empirical research on eyewitness testimony should be under-
taken. The Supreme Court should not be faulted for the criteria outlined in Neil v Biggers, however, because there was little eyewitness research from which to draw in 1972.

As outlined in this article, future considerations for revising the criteria set forth in Neil v Biggers must include a reworking of the eyewitness certainty question. Simply dropping eyewitness certainty as a criterion may not be sufficient; research shows that triers of fact naturally base a great deal of their judgments on eyewitness certainty (e.g., see Wells et al., 1981). Thus, in order to handle adequately the issue of eyewitness confidence, some cautions regarding the limited utility of eyewitness certainty may be in order. This may seem peculiar to the courts, but it can be as important to note factors that are not useful for judging eyewitness accuracy as it is to note factors that are useful. The Court should also consider some elaborations on the degree of attention criterion as outlined herein as well as on the delay criterion. The witness's opportunity to view the assailant might also be better explained to triers of fact, especially with regard to the matter of duration of exposure. The Court should acknowledge that "accuracy of prior description" really means agreement between the prior description and the identified suspect and that such agreement can exist for numerous reasons that do not necessarily imply accuracy. Finally, the Court should reestablish the concern with suggestiveness, possibly by including it as one of the central criteria, because of its strong influence on eyewitness accuracy.

Meanwhile, eyewitness researchers should continue to research features of eyewitness testimony that can yield results that are useful to the true needs of police investigators and courts (see Clifford, 1979 on this point). Although we have considerable knowledge to impart to the courts to help revise and refine their criteria for accepting and rejecting eyewitness accounts, there remain many areas in which we can improve on existing knowledge. Particularly salient in this regard is the fact that most of our current research tends to yield results that tell the courts what not to do (e.g., don't rely on eyewitness certainty, rely less on witness's prior descriptions of the assailant as indicants of accuracy, be reluctant to accept witnesses' estimations of short temporal durations). More research should be devoted to discovering things that the courts can do.

The majority of eyewitness research studies also make the common mistake of failing to consider the practical realities of how criminal justice personnel must obtain information. Consider the following hypothetical example. An eyewitness researcher stages thefts under one of two conditions. In one condition the confederate/thief looks directly at the eyewitness throughout the theft, whereas in the other condition he or she varies between a three-quarters profile and a direct profile. The latter condition is generally superior for subsequent facial recognition (Krouse, 1981). However, it could be misleading to conclude that real-world eyewitnesses who claim to have had a profile view of the criminal are more accurate than those witnesses who claim to have only had a frontal view. Specifically, the experimenter has obtained predictive utility from profile information because he or she knows the actual conditions of encounter. Courts, however, must obtain such information from the eyewitness's self-report. Our research, for example, indicates that eyewitnesses almost invariably report having viewed every possible profile (straight on, left three quarters, right three quarters) regardless of the actual witnessing conditions and that such reports have no relationship to eyewitness identification accuracy. Thus, profile information might be a diagnostic predictor of eyewitness accuracy when actual profile exposures are known by the predictor but a poor predictor when obtained via self-reports by the eyewitness. The general point here is that an eyewitness's self-report is often the only means available to courts for discerning the accuracy of the eyewitness; thus research should not only show that actual witnessing conditions can be used to predict eyewitness accuracy, but that self-reports of those conditions also have predictive value.

Eyewitness testimony could be one of the next major issues in which the Supreme Court looks to the psychological sciences for advice. It is important, therefore, that experimental research focus on the issues in a
manner that will prove useful to the Court. This article, which characterizes the Neil v Biggers criteria as a set of hypotheses that currently guide U.S. courts, is one framework around which useful empirical research can be built. By considering a framework of this type (i.e., one based on formal court criteria), eyewitness experiments can be designed that might have high forensic relevance.

Reference Notes

1 Brigham, J C, & Wolfskiel, M P Opinions of attorneys and law enforcement personnel on the accuracy of eyewitness identifications Unpublished manuscript, Florida State University, 1982
2 Wolfskiel, M P, & Brigham, J C The relationship between accuracy of prior description and facial recognition Unpublished manuscript. Florida State University, 1982
4 Fanselow, M S, & Buckhout, R Nonverbal cuing as a source of biasing information in eyewitness identification testing Unpublished manuscript, Brooklyn College, 1976

References

Arbuckle, T Y, & Katz, W A Structure of memory traces following semantic and nonsemantic orientation tasks in incidental learning Journal of Experimental Psychology Human Learning and Memory 1976, 2 362–369
Bahrick, H, Bahrick, P, & Wittlinger R Fifty years of memory for names and faces A cross-sectional approach Journal of Experimental Psychology: General 1975, 104, 54–75
Billig, M, & Millner, D A spade is a spade in the eyes of the law Psychology Today February 1976, pp 13–15
Block, E B The vindicators New York Alvin Redman, 1963
Borchard, E M Convicting the innocent Errors of criminal justice New Haven, Conn Yale University Press 1932
Clifford, B R Police as eyewitnesses New Society, April 1976, pp 176–177
Clifford, B R The relevance of psychological investigation to legal issues in testimony and identification Criminal Law Review 1979, 32, 153–163
Clifford, B R, & Richards, V J Comparison of recall by policemen and civilians under conditions of long and short durations of exposure Perceptual and Motor Skills, 1977, 45, 503–512
Craik, F I M, & Lockhart R S Levels of processing A framework for memory research Journal of Verbal Learning and Verbal Behavior 1972, 11, 671–684
Craik, F I M, & Tulving, E Depth of processing and the retention of words in episodic memory Journal of Experimental Psychology General, 1975, 104, 268–294
Davies, G M, Shepherd, J W, & Ellis, H D Similarity effects in face recognition American Journal of Psychology 1979, 92, 507–523
Deffenbacher, K Eyewitness accuracy and confidence Can we infer anything about their relationship? Law and Human Behavior 1980, 4, 243–260
Ebbinghaus, H Memory A contribution to experimental psychology New York Dover, 1964 (Originally published, 1885)
Egan, D, Pittner, M, & Goldstein, A G Eyewitness identification: Photographs vs live models Law and Human Behavior 1977, 1, 199–206
Egeth, H E, & McCloskey, M Expert testimony about eyewitness behavior Is it safe and effective? In G L Wells & E F Loftus (Eds) Eyewitness testimony Psy-
chological perspectives New York Cambridge University Press, 1983


Frankfurter, F The case of Sacco and Vanzetti New York Little, Brown, 1927

Gardner, E The court of last resort New York Little, Brown, 1952


Howells, T H A study of ability to recognize faces Journal of Abnormal and Social Psychology, 1938, 33, 124–127

Kahneinan, D Attention and effort Englewood Cliffs, N J Erlbaum, 1973


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


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


Loftus, E F Leading questions and the eyewitness report Cognitive Psychology, 1975, 7, 560–572


Loftus, E F Eyewitness testimony Cambridge, Mass Harvard University Press, 1979


Loftus, E F, & Greene, E Warning Even memory for faces may be contagious Law and Human Behavior, 1980, 4, 323–334


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


Malpass, R S, & Devine, P G Realism and eyewitness identification research Law and Human Behavior, 1980, 4, 347–358


Marshall, J Law and psychology in conflict New York Bobbs-Merrill, 1966


Munsterberg, H On the witness stand New York Doubleday, 1908

Murray, D M, & Wells, G L Does knowledge that a crime was staged affect eyewitness accuracy? Journal of Applied Social Psychology, 1982, 12, 42–53

Navon, D, & Gopher, D On the economy of the human processing system Psychological Review; 1979, 86, 214–255

Patterson, K E, & Baddeley, A D When face recognition fails Journal of Experimental Psychology Human Learning and Memory, 1977, 3, 406–407


Tickner, A, & Poulton, E Watching for people and actions Ergonomics, 1975, 18, 35–51


---

Received August 30, 1982
Revision received January 24, 1983

---

**Third Edition of the Publication Manual**


All manuscripts to be published in the 1984 volumes of APA’s journals will be copy edited according to the third edition of the *Manual* Therefore, manuscripts being prepared now should be prepared according to the third edition. Beginning in 1984, submitted manuscripts that depart significantly from third edition style will be returned to authors for correction.

The third edition of the *Publication Manual* is available for $12 for members of APA and $15 for nonmembers. Orders of $25 or less must be prepaid. A charge of $1.50 per order is required for shipping and handling. To order the third edition, write to the Order Department, APA, 1400 N. Uhle Street, Arlington, VA 22201.