Journal of Applied Psychology
1985, Vol. 70, No. 3, 547-555


0021-9010/85/$00.75

# Relationship Between Accuracy of Prior Description and Facial Recognition

Melissa Pigott and John C. Brigham
Florida State University

The present study empirically assessed the appropriateness of one of the Supreme Court's guidelines for the evaluation of eyewitness identification evidence. This guideline concerns the relationship between the accuracy of an eyewitness' description of a suspect and the witness' accuracy in his or her lineup decision. One hundred twenty college-aged male and female subjects were exposed briefly to a target person and were then required to describe his physical characteristics. Later they were asked to identify the target person from a photograph lineup in which he was or was not present. Subjects' certainty in their decision was also assessed. The results provided no support for the validity of the Supreme Court's guideline—there was no relationship between description and identification accuracy or between a subject's description and the characteristics of the person identified (rightly or wrongly) from the lineup. When all subjects who identified someone from the lineup were combined, a substantial relationship between confidence and accuracy was found. Theoretical issues concerning the effects of differences between target persons and between witnessing conditions are discussed.

The U.S. Supreme Court, in recognizing the limitations of eyewitness testimony, has imposed legal standards that define appropriate eyewitness identification techniques. In *Neil v. Biggers* (1972), the Court specified five conditions to be considered in the evaluation of identification evidence: (a) The opportunity of the witness to view the criminal at the time of the crime, (b) The length of time between the crime and the identification, (c) The level of certainty demonstrated by the witness at the identification, (d) The witness' degree of attention during the crime, and (e) The accuracy of the witness' prior description of the criminal. These five factors were later restated by the Court in *Manson v. Brathwaite* (1977). While the first four of these factors have been the subject of considerable empirical investigation (see Wells & Murray, 1983, for a review of findings on the *Neil v. Biggers* criteria), empirical data are practically nonexistent with regard to the fifth factor, accuracy of prior descriptions. This last factor was the major focus of the present study.

Although verbal descriptions of suspects are an important part of police work, only a few empirical studies have attempted to investigate the relationship between verbal descriptions and facial recognition. Goldstein and Chance (1971) found that the ease with which a face was labeled was not related to accuracy of recognition. A further study (Chance & Goldstein, 1976) showed that subjects who were instructed to describe something about a face that would help them to recognize it later were only slightly more accurate in recognition performance than subjects who either generated word associates or only looked at the faces. Because verbalizations only slightly improved recognition accuracy, the authors concluded that there is only a weak involvement of verbal encoding in recognition memory for faces. Studies by Malpass, Lavigueur, and Weldon (1973) and

This research was supported by National Institute of Justice Grant (82-IJ-CX-0012) to John C. Brigham. Special gratitude goes to Rodger Pearson and Don Fann for their participation as confederates and for their help during all phases of the study. Thanks are also due to Sgt. Bill Taylor and Lt. Ted Rivenbark of the Florida State University Police Detective Unit for their assistance in lineup construction and in familiarizing the experimenters with the police procedures involved in obtaining eyewitness identifications and to Darryl Bruce, Robert Bothwell, Gary Wells, and an anonymous reviewer for comments on an earlier draft of this article.

Requests for reprints should be sent to John C. Brigham, Department of Psychology, Florida State University, Tallahassee, Florida 32306.

by Goldstein, Johnson, and Chance (1979) also failed to find a relationship between facial recognition and verbal encoding.

Research findings from other paradigms also suggest that people's ability to recall something seen previously will not be strongly related to their ability to identify the same object on a recognition test. A large body of research (see Flexser & Tulving, 1978; 1982, for reviews of these studies) demonstrates that persons who can recall a word when provided a retrieval cue often cannot recognize the word on a recognition test. This general finding, based on several variations in paradigm (e.g., see Tulving & Thomson, 1973), has been termed *retrieval independence,* meaning that cues that are present on a recognition task are uncorrelated with those on a recall test. These findings would suggest that, counter to the Supreme Court's assertion, accuracy of prior descriptions (recall) is not likely to be related to ability to identify someone from a lineup (recognition). However, this has never been assessed in a forensically relevant situation.

The present study tested the Court's assumption by simulating the procedures involved when a witness describes and later identifies an offender. In addition, several other factors were investigated for their effect on the accuracy of subjects' descriptions and identifications. Two variables, degree of subjects' attention and certainty, were of particular interest as they are also included in the Supreme Court's guidelines.

To see whether the relationship between accuracy of prior description and accuracy of identification differed when levels of attention were different, a manipulation intended to create two levels of attention by varying depth of processing was employed. Studies of memory for pictures have suggested that relatively deep processing is caused by focusing subjects' attention on the personality traits of a pictured individual, whereas shallow processing is engendered by focusing attention on a physical trait of a pictured individual (e.g., Bower & Karlin, 1974; Chance & Goldstein, 1981; Winograd, 1976, 1978, 1981). It was predicted that both identification accuracy and description accuracy would be greater in the high attention/deep processing conditions than in the low attention/shallow processing conditions, but that identification accuracy and description accuracy would not be related within or across the two attentional levels.

The type of lineup was also varied. The target person was either present or absent from the photo lineup that subjects viewed. This was done to determine whether subjects would choose the person who best fit their description even when the actual target person was not present in the lineup. Subjects' degree of self-monitoring was also assessed because some previous research suggests that this personality variable may be related to eyewitness accuracy (Hosch, Leippe, Marchioni, & Cooper 1984; Hosch & Platz, 1984). It was hypothesized, in line with the Hosch et al. (1984) findings, that subjects high in self-monitoring would be more accurate in their descriptions and identifications than subjects low in self-monitoring.

A fifth main analysis investigated the relationship of subjects' certainty about their identification and their actual accuracy. Although the research data on this question are mixed (see Deffenbacher, 1980; Leippe, 1980; Wells & Murray, 1984 for reviews) it was predicted that a positive relationship would be observed between identification accuracy and certainty for persons who chose someone from the lineup (cf. Malpass & Devine, 1981a). A positive relationship for choosers indicates that persons who chose someone from the lineup and were confident in the accuracy of their decision are likely to be accurate. In contrast, the relationship between identification accuracy and certainty for persons not making a lineup choice was not expected to be significant.

Although the U.S. Supreme Court used the word "accuracy" when writing of the criterion to which a witness' description should be compared, some have argued (Wells & Murray, 1983) that the court was really addressing the congruence between the witness' description and the characteristics of the person identified from the lineup. This interpretation would suggest a different type of analysis, focusing on the correspondence between the subject's description and the characteristics of the lineup member selected,

whether target or foil.[1] This congruency analysis was carried out in addition to the accuracy analysis.

## Method

### Overview of Design

The present study involved a 2 (attention/depth of processing: deep or shallow) by 2 (type of photo lineup: target present or target absent) by 2 (target person) factorial design, with all factors serving as between-subjects factors. Fifteen subjects were in each cell. Two target persons were employed to increase the generalizability of the findings. Individual subjects were randomly assigned to one of the attention/depth of processing conditions and the type of lineup viewed was randomly determined for each subject. Dependent measures were subjects' description accuracies, identification accuracy, self-monitoring scores, and identification certainty scores.

### Subjects

Subjects were 120 male and female undergraduates enrolled in introductory psychology classes at Florida State University. All subjects received course research credit for their participation.

### Instruments

A checklist of the type frequently used by police was administered to the subjects after the departure of the target person. This checklist required subjects to describe the confederate's physical characteristics.[2] A self-monitoring questionnaire (Snyder, 1974, 1979) was administered to all subjects. This scale consists of 25 true–false items designed to measure social interaction styles related to the observation and control of one's expressive behavior and self-presentation. Persons scoring high on the scale are those who are concerned with social appropriateness; that is, they are more concerned than others about behaving correctly in social situations.

Also used were four sets of six-photograph lineups constructed with the aid of the F.S.U. Campus Police Department. All photographs were Polaroid color prints (7.25 cm × 9.5 cm) of lineup members' head and shoulders. Two lineups, one target present and one target absent, were constructed for each white male confederate. The foil photos were chosen by the F.S.U. Police from a group of 25 photos of college-age white males.

### Development of Lineups

In addition to adopting the standards of the F.S.U. Police for lineup construction, mean similarity ratings were compiled for each potential foil photo. The similarity ratings, based on the ratings of 60 pilot subjects (30 for each target person), measured each photo's degree of similarity to the target person on a 7-point scale. All of the foil photos chosen for the lineups had mean similarity ratings between 5 ("somewhat similar") and 7 ("very similar"). In the target-absent lineups, the photo receiving the highest similarity rating was substituted for the target person (highest mean similarity ratings were 6.7 and 6.4), with the other foils remaining the same as in the target-present lineup.

Two other pilot studies were performed to obtain "accurate" descriptions of the targets and foils with which to compare the subjects' descriptions. Descriptions of the target persons were obtained from 50 college-student raters (25 for each target person) who viewed them for as long as they wished while completing the checklist. Pilot subjects' responses were tabulated into means and standard deviations for the age, height, and weight variables. Modes were computed for the discrete variables. "Accurate" criterion descriptions for each foil photo in Target A's lineup and 1 additional lineup (described below) were obtained from 96 pilot subjects (16 per photo) in a similar manner.

### Procedure

Groups of 3 to 6 subjects were told that they were participating in a study concerning expressive behavior and self-presentation. Subjects were seated the same distance (4.5 M) from the front of the experimental room and given a booklet containing instructions and the dependent measures. The instructions stated that in a few minutes a person would enter the room. Subjects were told that they would later be asked to estimate how honest he was (for subjects in the high-attention/deep-processing condition) or how tall he was (for subjects in the low-attention/shallow-processing condition). One of the target persons then walked into the room, stood facing the subjects for 15 s, and turned around and left the room.

After the target person's departure, subjects completed the description checklist and the self-monitoring questionnaire. Following this, subjects individually viewed one of the six-photograph lineups (target present or target absent) for the confederate they had seen. Instructions for the lineup phase of the experiment were as follows: "A photo of the person you just saw and described may or may not be included in this set of photos. Please look carefully at all of the pictures and indicate whether you see a photo of the person you just saw and described." Whether or not they chose a photo, all subjects rated their confidence in the accuracy of their decision on a scale ranging from 1 (uncertain that one's choice was correct) to 7 (certain that one's choice was correct).[3] After this, subjects were completely debriefed.

---

[1] We are indebted to Gary Wells for suggesting this additional interpretation and analysis.

[2] The checklist included: race, sex, ethnicity, age, height, weight, body build, hair color, hair length, hairstyle, skin color, eye color, eye shape, eyebrows, face shape, facial hair, nose shape, lip shape, ears, forehead, distinguishing marks (such as scars or tattoos), and clothing. The reliability coefficient of the checklist is .60.

[3] The measurement of a witnesses' confidence becomes problematical when the witness fails to choose anyone from the lineup. Ideally, one might be interested in the witnesses' confidence in the *decision* made (i.e., the decision that Person X from the lineup is the target

### *Description Accuracy and Description Congruency Scores*

The accuracy of subjects' descriptions was calculated by comparing their descriptions with those obtained from the raters. As no precedent has been set for the scoring of person's physical descriptions, four different measures of description accuracy were employed.[4] Subsequent analyses determined these measures to be related (*r*s from .55 to .65, all *p*s < .001). Because overlapping data were used in the calculation of the description accuracy measures, this lack of independence was expected.

In addition to the description accuracy scores, a description congruency score was calculated for all subjects who chose someone from the photo lineup. The congruency analysis involved a comparison of a subject's completed description checklist with a description of the person (target *or* foil) chosen from the lineup.

## Results

### *Identification Accuracy Rates*

Identification accuracy scores were obtained by calculating the proportion of hits for subjects viewing the target-present lineups or the proportion of correct rejections for subjects viewing the target-absent lineups. Identification accuracy rates are presented in Table 1. Overall identification accuracy was 70.83%. This accuracy rate was not affected by type of lineup, target-present or target-absent (70.0% vs. 71.7%, respectively; *z* for proportions = .202, *ns*). Contrary to predictions, similar accuracy rates were obtained regardless of the attention manipulation (75.0% and 66.8% for high and low attention,

Table 1
*Identification Accuracy Rates*

| | Target A attention | | Target B attention | | Mean accuracy for lineups |
|---|---|---|---|---|---|
| | Low | High | Low | High | |
| Target-present lineup | 86.7 | 93.3 | 53.3 | 46.7 | 70.0 |
| Target-absent lineup | 66.7 | 73.3 | 60.0 | 86.7 | 71.7 |
| Mean accuracy for targets | 80.0 | | 61.7 | | |
| Mean accuracy for attention | | | | | |
| Low | | | | | 66.8 |
| High | | | | | 75.0 |
| Overall accuracy | | | | | 70.8 |

*Note.* Figures are percentages. $N$ = 15 per cell (total $n$ = 120).

(person *or* that no one in the lineup is the target person). Malpass and Devine (1981a) asked their subjects how confident they were that the vandal was in the lineup and also calculated a confidence index based on subjects' confidence ratings for each lineup photo. In the present study, witnesses were asked how confident they were that they had chosen the right photo. In both the Malpass and Devine (1981a) study and the present study, witnesses who picked someone out of a lineup had significantly higher confidence scores than those who chose no one from the lineup (*M*s = 5.27 and 3.00 in the present study). In fact, the overall point-biserial correlation between choosing and confidence in the current study was $r_{bp}$ = .91 ($p < .001$). The fact that the concepts of choosing and confidence are so closely related in witnesses' minds, coupled with the fact that no-choice situations are of lesser ecological validity for criminal identifications, makes the calculation of confidence–accuracy correlations within target-absent conditions difficult to interpret (cf. Malpass & Devine, 1981a). Combining the data from witnesses who choose and who do not choose for one grand confidence–accuracy correlation is inappropriate, because the confidence ratings of no-choice subjects will be lower and are likely not to have the same meaning as confidence ratings by subjects who made a choice. Had this procedure been followed in the current study, the resulting correlation would have been $r$ = .03, leading to the (incorrect) conclusion that accuracy and confidence were not related.

[4] For the first description-accuracy measure, responses for each description were divided into categories and scored as follows: 4 = *very accurate* (matching actual data or raters' modal response); 3 = *slightly accurate* (a small difference from actual data or raters' modal response); 2 = *slightly inaccurate* (a moderate difference from actual data or raters' modal response); and 1 = *very inaccurate* (a large difference from actual data or raters' modal response). Descriptions involving the actual measured characteristics of the confederates (age, height, and weight) were scored in terms of their degree of standard deviation from the correct response. Factors not mentioned by a particular subject were not included in the calculation of his or her mean description accuracy score.

The second description accuracy score was derived by multiplying a subject's mean accuracy score by the number of traits that he or she attempted to describe. The third measure of description accuracy was similar to the first except that the third measure treated subjects' nonresponses as zeroes, such that an individual's mean score would be lowered by omitting any feature description. The last description measure treated responses as either right or wrong. For a subject's response to be correct, it had to exactly match the rater's mean or modal response for that particular characteristic. Correct responses were scored at +1. All other responses were considered incorrect and were scored as 0.

respectively; $z$ for proportions = 1.19, *ns*), suggesting that the attention manipulation had not been successful. However, there was a significant difference in identification accuracies between target persons (80% and 62%, $z$ for proportions = 2.14, $p < .05$).

*Analyses of Identification Accuracy*

A 2 × 2 × 2 multiple contingency analysis (Winer, 1971) was performed on subjects' identification accuracy scores. Of the seven calculations that were performed, two were statistically significant. The effect of target person on accuracy was significant, $\chi^2(1) = 4.88$, $p < .05$; Target A was identified more accurately overall than Target B (see Table 1). The interactive effect of target person and type of lineup on accuracy ($\chi^2 = 6.07$, $p < .05$) was also significant. For Target A, accuracy was greater for target-present lineups (90.0%) than for target-absent lineups (70.0%), whereas accuracy for Target B was less in target-present lineups (50.0%) than for target-absent lineups (73.3%). A 2 × 2 × 2 analysis of variance (ANOVA) replicated the above findings, $F(1, 61) = 5.20$, $p < .025$, $\text{eta}^2 = .04$, for the effect of target person on accuracy, $F(1, 61) = 7.26$; $p < .008$, $\text{eta}^2 = .06$, for the interaction of target person and type of lineup.

*Analyses of Description Accuracy*

A 2 × 2 × 2 factorial ANOVA (Attention × Type of Lineup × Target Person) was performed on each of the four description accuracy scores. There was a consistent significant main effect for target person across all four measures of description accuracy, with Target A described more accurately, $F$'s(1, 119) ranged from 3.98 to 12.17, $p$s from .048 to .001, $\text{eta}^2$s from .03 to .09. There was no interactive effect, however, of target person and type of lineup on accuracy. Significance was not obtained in any of the four ANOVAs for the attention or lineup variables or for the interactions among variables.

*Identification Accuracy–Description Accuracy Relationship*

Point-biserial correlations were calculated between each of the four measures of description accuracy and the identification accuracy scores and between congruence and identification accuracy, within each condition and across all conditions. Correlations between the four description accuracy scores (and congruence scores) and identification accuracy were separately assessed for persons making and not making a lineup choice. In the current study, 75.0% ($N = 45$) of witnesses exposed to the target-present lineup and 28.3% ($N = 17$) of the witnesses exposed to the target-absent lineup picked someone from the lineup. Only 6 of the 135 correlations (all combinations of target, target presence/absence, and attention, analyzed separately and in all combinations) for choosers reached significance, and none of the five correlations across all conditions were significant, $r$s(60) from −.16 to .12, *ns*. None of the 135 correlations for nonchoosers reached significance, overall $r$s(56) from −.11 to .04, *ns*. Correlations were also calculated between the mean identification accuracy rates in each condition (see Table 1) and the four corresponding description accuracy scores to see whether conditions that produced greater identification accuracy also yielded better description accuracy. None of the four correlations reached significance, $r(6)$ from −.39 to −.44, *ns*.

Point-biserial correlations were also calculated between identification accuracy and description congruence scores for the conditions involving Target A. There were no significant results, overall $r(35) = .16$, *ns*. These results replicate those discussed above, thus providing additional support for the finding of independence between description accuracy and identification accuracy.[5]

*Confidence–Identification Accuracy Relationship*

Legal and criminal justice personnel have been most concerned with the situation when

[5] Target B's photo lineup was no longer available when this additional analysis was undertaken. To gather further evidence on this point, we also analyzed the congruence–accuracy relationship for two additional lineups that had been employed in a study similar to the present one (Brigham & Pigott, 1983). The second study also used Target A; the findings paralleled those for the present study: No relationship was found between congruence and identification accuracy, $r(30) = .01$, *ns*, for Target A's lineup; $r(34) = -.25$, *ns*, for the other lineup.

a witness identifies a suspect (makes a choice). The witness's expressed level of confidence may then be used in the courtroom to buttress the identification testimony. The most relevant comparison regarding this situation in the present study is the confidence–accuracy relationship for all witnesses who made a choice in both target-absent and target-present conditions (because, in an actual eyewitness situation, it is a matter of dispute whether the lineup shown the witness was target-absent or target-present). For witnesses who picked somebody from the lineup, the accuracy–confidence correlation was $r(60) = .512$ ($p < .001$). For purposes of comparison, within the target-present conditions the confidence-accuracy correlation for choosers was $r(43) = .413$, $p < .01$. Table 2 presents the confidence-accuracy correlations for all witnesses who made a choice, according to which target person they observed and level of attention. The confidence–accuracy correlation for the target for whom the accuracy rate was significantly lower (Target B) was significantly higher than the confidence–accuracy correlation for Target A ($z = 2.06$, $p < .05$). Confidence–accuracy correlations were also calculated for subjects who did not make a lineup choice. The overall correlation for nonchoosers was not significant, $r(56) = .09$, *ns*.

*Self-Monitoring and Accuracy*

Correlations between self-monitoring and the two accuracy measures (description and identification) were also calculated both within and across conditions. Because statistical significance was obtained in only two of these 45 comparisons, it seems most appropriate to treat the two significant findings as chance occurrences. Correlations between self-monitoring and description congruence likewise yielded no significant results.

Table 2
*Confidence–Identification Accuracy Correlations for All Witnesses Choosing Someone From the Lineup, Target-Present and Target-Absent Conditions Combined*

| | Attention | | |
|---|---|---|---|
| Target | Low | High | Combined attention |
| Target A | .233[a] | .468*[b] | .329* |
| Target B | .603**[c] | .893***[d] | .716*** |
| Combined targets | .406** | .637*** | |

*Note.* Overall $r = .512$**.
[a] $N = 19$. [b] $N = 18$. [c] $N = 16$. [d] $N = 9$.
* $p < .05$. ** $p < .01$. *** $p < .001$.

## Discussion

The present study is the first empirical test in a forensically relevant situation of the U.S. Supreme Court's guideline regarding accuracy of prior descriptions. The results provide no support for the guideline. There was no relationship between subjects' accuracy in describing the target person and the accuracy with which they recognized him in a photograph lineup. Similarly, there was no relationship between description congruence and identification accuracy. Contrary to the Supreme Court's guideline, it cannot be assumed that persons who are accurate in describing another person will also be accurate in recognizing that person.

The observed results are in agreement with those of earlier laboratory face-recognition studies (e.g., Chance & Goldstein, 1976; Goldstein & Chance, 1971; Goldstein et al., 1979; Malpass et al., 1973) in finding no significant relationship between facial recognition and description accuracy. The results are also similar to those of experiments showing recognition/recall independence (e.g., see Flexser & Tulving, 1978; Tulving & Thomson, 1973). Applying results from these memory experiments to the present study, one could argue that the retrieval cues that were available to subjects during completion of the description checklist (a cued recall task) did not assist them in identifying the target from the photo lineup (a recognition test). Thus, a subject who accurately described the target person could not, as a consequence, be expected to more accurately recognize him later than someone whose initial description was relatively inaccurate.

The findings indicate that the depth of processing manipulation that has been effec-

tive for pictorial material was not effective for our live stimulus persons, at least in terms of predicted differences in accuracy. The more complicated procedures employed in the present study, where subjects saw the stimulus person, then attempted to describe him, answered a self-monitoring questionnaire, and attempted an identification from a photo lineup are quite different from typical picture recognition studies. In addition, seeing a live stimulus person may have been so interesting and involving that differences in orienting instructions became largely irrelevant. Hence the adaption of a widely used manipulation from one paradigm to another proved problematical.

This is particularly unfortunate because although we assessed the identification–description relationship within the general level of attention engendered by the experimental situation, we are not able to make a strong between-conditions comparison. Our data show that good describers were not good identifiers in this general situation, but they cannot address the related between–conditions question of whether some witnessing conditions (e.g., very high attention, very low attention) may affect *both* identification and description accuracy in the same direction. If this were the case, then the description–identification correlation across conditions could be significant even when each within–condition correlation was near zero. Hence, it would be premature, based on the present data, to conclude that there will not be any correlational relationship between identification and description accuracy across varying levels of important factors (such as, perhaps, attention or arousal). What we can conclude is that at a given level of attention there was no evidence of a positive relationship between description and identification accuracy.

The finding that Target A was identified significantly more accurately (90%) than Target B (50%) in target-present lineups, whereas no difference occurred for target-absent lineups, underscores the importance of the nature of the target stimulus. One might suspect that Target A was more distinctive in appearance than Target B, a supposition that is supported by the finding that description accuracy for Target A was significantly higher in the current study, and distinctiveness ratings gathered for these two targets in another study were higher for Target A than for B (Brigham & Pigott, 1983). Moreover, calculations of the functional size (Wells, Leippe, & Ostrom, 1979) of the target-present lineups in the current study indicated (based on 110 additional raters) that Target A's lineup had a considerably smaller functional size (8.00) than Target B's lineup (20.00), meaning that A's lineup was "easier" than B's, presumably because of A's distinctiveness. These differences occurred despite our considerable efforts to create fair and equivalent lineups for A and B based on similarity ratings and our Police Department consultant.

Despite these target-person differences, the (absence of a) relationship between description and identification accuracy was similar for the two targets. The confidence–accuracy correlation for choosers was significantly lower for Target A than for B, a finding that merits further investigation. The substantially higher accuracy rate among choosers for Target A may have depressed the point-biserial correlation because of restricted variance. These findings suggest that the widely noted mixed nature of findings on confidence and accuracy, as reviewed by Deffenbacher (1980), Leippe (1980), and Wells and Murray (1984), may stem partly from differences in the stimulus persons used.

The substantial overall correlation between confidence and accuracy for subjects who made a choice is consistent with some recent research (e.g., Brigham, Maass, Snyder, & Spaulding, 1982; Fleet, Brigham, & Bothwell, 1984; Maass & Brigham, 1982; Malpass & Devine, 1981b) that has found confidence–accuracy correlations of this magnitude. Other researchers, however, have found little or no relationship between these variables (e.g., Brown, Deffenbacher, & Sturgill, 1977; Leippe, Wells, & Ostrom, 1978; Wells, Ferguson, & Lindsay, 1981). A recent meta-analysis of 26 studies investigating the confidence–accuracy relationship (Bothwell, Brigham, & Deffenbacher, 1985) has shown the results to be extremely variable (C.I. [95%] for $r = -.03, .54$). The present overall confidence–accuracy correlation (.51) is near the upper end of these confidence limits.

In conclusion, we found no empirical support for the assumption that an eyewitness who accurately describes a perpetrator of a crime will be more accurate in identifying him or her than a witness whose initial description is less accurate. Therefore, the validity of the guideline imposed by the Supreme Court can be questioned. Further research that effectively varies situational factors such as attention or level of arousal will be valuable in assessing whether significant covariation between description and identification accuracy may occur across levels of these variables. The question that seems of most direct legal relevance is as follows: In a given situation, should one place more credence in an identification made by someone who has given an accurate description than someone who has given an inaccurate description, as suggested in *Neil vs. Biggers?* The answer appears to be no. The answer to the broader question of some theoretical interest, whether witnessing conditions can affect both description accuracy and identification accuracy, awaits further research.

## References


Barkowitz, P., & Brigham, J. C. (1982). Recognition of faces: Own-race bias, incentive, and time delay. *Journal of Applied Social Psychology, 12,* 255–268.

Bothwell, R. K., Brigham, J. C., & Deffenbacher, K. A. (1985). *A meta-analysis of the influence of confidence and arousal on eyewitness accuracy.* Unpublished manuscript, Florida State University.

Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology, 103,* 751–757.

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

Brigham, J. C., & Pigott, M. P. (1983). *Comparison of description and identification accuracy under different attentional levels.* Unpublished manuscript, Florida State University.

Brown, E., Deffenbacher, K., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology, 62,* 311–318.

Chance, J., & Goldstein, A. (1976). Recognition of faces and verbal labels. *Bulletin of the Psychonomic Society, 7,* 384–387.

Chance, J., & Goldstein, A. (1981). Depth of processing in response to own- and other-race faces. *Personality and Social Psychology Bulletin, 7,* 475–480.

Deffenbacher, K. (1980). Eyewitness accuracy and confidence. *Law and Human Behavior, 4,* 243–260.

Ellis, H. D. (1975). Recognizing faces. *British Journal of Psychology, 6,* 409–426.

Fleet, M. L., Brigham, J. C., & Bothwell, R. K. (1984). *The effects of choosing on the confidence–accuracy relationship in eyewitness identifications.* Unpublished manuscript, Florida State University.

Flexser, A. J., & Tulving, E. (1978). Retrieval independence in recognition and recall. *Psychological Review, 85,* 153–171.

Flexser, A. J., & Tulving, E. (1982). Priming and recognition failure. *Journal of Verbal Learning and Verbal Behavior, 21,* 237–248.

Goldstein, A., & Chance, J. (1971). Visual recognition memory for complex configurations. *Perception and Psychophysics, 9,* 237–241.

Goldstein, A. G., Johnson, K. S., & Chance, J. (1979). Does fluency of face description imply superior face recognition? *Bulletin of the Psychnomic Society, 13*(1), 15–18.

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

Hosch, H. M., & Platz, J. J. (1984). Self-monitoring and eyewitness accuracy. *Personality and Social Psychology Bulletin, 10,* 289–292.

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

Leippe, M. R., Wells, G. L., & Ostrom, T. M. (1978). Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology, 63,* 345–351.

Maass, A., & Brigham, J. C. (1982). Eyewitness identifications: The role of encoding specificity and attention. *Personality and Social Psychology Bulletin, 8,* 54–59.

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

Malpass, R. S., & Devine, P. G. (1981b). Guided memory in eyewitness identification. *Journal of Applied Psychology, 66,* 343–350.

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face recognition. *Perception and Psychophysics, 14,* 285–292.

*Manson v. Brathwaite,* 432 U.S. 98, 97 S. Ct. 2243, 53 L. Ed. 2d 140 (1977).

*Neil v. Biggers,* 409 U.S. 188, 93 S. Ct. 375, 34 L. Ed. 2d 401 (1972).

Shepherd, J. W., Davies, G. M., & Ellis, H. D. (1980). *Identification after delay.* Final Report for Grant RES 522/4/1 from the Home Office Research Unit, University of Aberdeen, Scotland.

Snyder, M. The self-monitoring of expressive behavior. (1974). *Journal of Personality and Social Psychology, 30,* 56–537.

Snyder, M. Self-monitoring processes. (1979). In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12), (pp. 85–128). New York: Academic Press.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

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

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

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

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

Winograd, E. (1976). Recognition memory for faces following nine different judgements. *Bulletin of the Psychonomic Society, 8,* 419–421.

Winograd, E. (1978). Encoding operations which facilitate memory for faces across the life span. In M. N. Gruenberg, P. E. Morris, & R. N. Sikes (Eds.), *Practical aspects of memory* (pp. 255–262). New York: Academic Press.

Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 181–190.

Winer, B. J. (1971). *Statistical principles in experimental design.* New York: McGraw-Hill.

Received March 9, 1984
Revision received December 3, 1984 ■

---

## Correction to Doverspike and Barrett

In the article "An Internal Bias Analysis of a Job Evaluation Instrument" by Dennis Doverspike and Gerald V. Barrett (*Journal of Applied Psychology,* 1984, Vol. 69, No. 4, pp. 648–662), there were errors in Table 2. For the female sex-typed jobs, the correlations between Scale 6, Physical Effort, and Scales 11–15 should be as follows: .25, .16, .54, .67, and −.26. In addition, again for the female sex-typed jobs, the correlations with Total Points were as follows: $r = .39$ for Physical Effort (6), $r = .68$ for Safety of Others (10), $r = .81$ for Counseling (11), $r = .81$ for Negotiating (12), $r = .38$ for Surroundings (15), $r = .49$ for Hazards (14), and $r = -.53$ for Monotony. Only the change for the correlation between Physical Effort and Total Points changes the results: The female sex-typed correlation is now significantly higher than the male sex-typed correlation.