Journal of Personality and Social Psychology
1978, Vol. 36, No. 12, 1546-1557

Applied Eyewitness-Testimony Research: System Variables and Estimator Variables

Gary L. Wells
University of Alberta, Edmonton, Canada

A distinction is proposed between two types of applied eyewitness-testimony research: *System-variable research* investigates variables that are manipulable in actual criminal cases (e.g., the structure of a lineup) and, thus, has the potential for reducing the inaccuracies of eyewitnesses; *estimator-variable research*, however, investigates variables that cannot be controlled in actual criminal cases (e.g., characteristics of the witness) and, thus, can only be used in the courtroom to augment or discount the credibility of eyewitnesses. System variables and estimator variables are contrasted with respect to their relative potential for positive contribution to criminal justice, and it is concluded that system-variable research may prove more fruitful than estimator-variable research. It is also argued that several methodological biases may be exacerbating the rate of misidentifications in staged-crime paradigms.

Clearly, some eyewitness-testimony research can be thought of as concerning basic, theoretical issues in perception, learning, and memory (see Loftus & Palmer, 1974, for a possible example). However, most research that is *specifically* directed at eyewitness testimony either explicitly states its concern as "applications to criminal justice" or implies such a concern by operationalizing variables in a "real-world" fashion. The concern in the current article is with this latter type —applied eyewitness-testimony research.

Despite evidence of psychological interest in applied eyewitness-testimony research that dates back to the early part of the 20th century (e.g., Whipple, 1909), the goal of such research has not been well defined. Generally, researchers allude to a concern with or application to criminal justice. However, in reviewing the eyewitness literature, one finds no statement regarding what is meant by such phrases as "applications to criminal justice," "concerns with the concept of criminal justice," and so on. In order to have a working definition, this article will assume that the goal of applied eyewitness-testimony research is to generate scientific knowledge that will maximize the chances that a guilty defendant will be justly convicted while minimizing the chances that an innocent defendant will be mistakenly convicted.[1] In addition, the term *criminal justice system* will refer to any governmental policy or practice that potentially affects criminal justice as defined above.

To the extent that eyewitness-testimony research has been directed toward applica-

The author would like to thank Anthony N. Doob, Tamara Ferguson, Anthony G. Greenwald, R. C. L. Lindsay, and E. F. Loftus for comments throughout the preparation 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] This article will assume that convicting an innocent defendant (false alarm) is no more of an injustice than is releasing a guilty defendant (miss). While this may seem to contradict the views of many social-action groups concerned with criminal justice, any differential weighting of these two types of error might restrict researchers' focus on the broader issues and/or inappropriately affect how researchers operationalize eyewitness variables. On the other hand, differential weighting might be defended by the often-overlooked fact that for "false alarms," the falsely accused suffers *and* the true criminal is still at large, whereas for "misses," only the latter factor is operative. In addition, a "miss" is not considered a vindication of the criminal suspect.

tions to criminal justice, the research has not had a great impact (Buckhout, Note 1; Hastie, Loftus, Penrod, & Winkler, Note 2).[2] Buckhout (Note 1) accounts for this weakness by characterizing the criminal justice system as reactionary and closed minded. Yet, the reactions of the criminal justice system may be quite understandable. For example, in Buckhout's (1974) oft-cited *Scientific American* article, Buckhout operates on the explicitly stated supposition that "eyewitness testimony is unreliable" (p. 23). Such blanket discountings of eyewitness testimony, even if they are correct, are bound to be greeted with negative reaction by the criminal justice system. Clearly, the criminal justice system will never eliminate eyewitness testimony altogether. It can be noted that (a) the initial report from a victim that a crime has occurred is in itself a testimony; (b) police officials' statements that they received a report of the crime is a testimony; and (c) to discount all eyewitness testimony is to discount any possibility of the defendant's having an alibi, thereby placing innocent defendants in a precarious situation.

Thus, this article begins with the position that the total *elimination* of eyewitness testimony is neither a desirable nor a feasible approach to improving criminal justice. In addition, it will later be argued that such a position is not necessarily a scientifically sound conclusion to be drawn from the eyewitness literature.

Hastie, Loftus, Penrod, and Winkler (Note 2) suggest that court reluctance to utilize scientific psychology in the courtroom may stem from the lack of an adequate published summary of relevant laboratory and field research. In line with this idea, the current author notes that there has not been a complete review (either applied or theoretical) of the eyewitness-testimony literature in an American Psychological Association journal since Whipple's *Psychological Bulletin* article in 1918. But the field of applied eyewitness-testimony research needs more than a review; it needs criticism and a collimated line of focus. Researchers of basic theory have maintained direction and focus by precisely demonstrating how their operations apply to a *theory*. Applied eyewitness-testimony researchers, however, have generally been somewhat unclear as to how their operations apply to criminal justice. Although it is the responsibility of individual researchers to make this link, it might prove fruitful to have a general category system that relates eyewitness variables to criminal justice concepts.

Investigators concerned with theoretical analyses of memory processes use a tripartite system that includes acquisition, retention, and retrieval phases. While it is possible to adopt this categorization schema in summarizing eyewitness research (see Hastie et al., Note 2, for an excellent example), it is not a categorization system that is inherently related to *criminal justice applications*. For purposes of criminal justice applications, a categorization of eyewitness research should relate eyewitness-research variables to their domain of utility for criminal justice. In the next section, a two-category system is proposed that might help eyewitness researchers define the applications of their research. In addition, a sampling of relevant research in each category as it relates to criminal justice is provided.

The reader is cautioned against assuming that the placement of a given piece of research into one category or the other category is, in itself, a comment on the *general* utility of that research. This article focuses only on *applied* utility, and the categorization system is designed with that in mind. There is no intention to "pigeonhole" research or discredit studies because they lack applied utility for criminal justice. Indeed, some of the research was not designed or billed by the investigators as applied eyewitness research. The purpose is to introduce a categorization system that may prove useful to researchers in designing their research

[2] This is not to say that the psychological literature on eyewitness testimony has been ignored by law journals or lawmakers. Levine and Tapp's (1973) article in the *Pennsylvania Law Review* outlines the psychological literature's relevance for the criminal identification process. This was recently updated by Woocher (1977) in the *Stanford Law Review*. A recent government-solicited report (Devlin, 1976) has also shown sensitivity to the apparent inadequacy of eyewitness identifications.

operations so as to increase the applied utility of their results if, in fact, they are interested in applications to criminal justice. Specific research examples from the literature are included primarily as an exemplary aid for the categorization system. The next section briefly introduces the categorization system.

Estimator and System Variables in Eyewitness Research

Fallible eyewitness testimony threatens the concept of criminal justice, and there are two types of research directed at this issue. The first type investigates variables that affect eyewitness accuracy but are not under the control of the criminal justice system. That is, although these variables may be manipulable in research, they cannot be controlled in actual criminal cases. Such variables will be termed *estimator variables* because, in actual crimes, one can at best only *estimate* the role of such factors. Independent control over estimator variables is, for all practical purposes, impossible for actual crimes. For example, moderately severe crimes may produce greater witness accuracy than very severe crimes (Johnson & Scott, Note 3), but the criminal justice system cannot directly control the severity of crimes so as to produce less fallible eyewitness accounts. The criminal justice system could, however, use such knowledge to *estimate,* post hoc, the likely accuracy of a witness.

The second type of eyewitness research investigates variables that are (or potentially can be) under the direct control of the criminal justice system. These variables will be termed *system variables* because of their relevance for application to change in the criminal justice system. An example of a system variable is the length of time between the initial criminal event and subsequent testimony. Evidence that this variable affects eyewitness-testimony performance (Shepard, 1967) might be directly used to advocate changes in the scheduling of police interrogations.

A Closer Examination of the Estimator–Variable Literature

There are several ways in which estimator-variable research might contribute to aiding criminal justice. One possibility is to assess experimentally the conditions that influence accuracy, "plug in" these conditions for an actual criminal case, and make a professional estimate regarding how likely it is that the witness(es) could give accurate or inaccurate testimony under such conditions. Another possibility is for the psychologist to appear in court as an expert witness and draw upon estimator-variable research to caution jurors and judges against the acceptance of certain types of testimony. There are other approaches, such as the *general* statement approach, to wit: "eyewitness testimony is unreliable." But, whatever the approach, estimator-variable research has a natural restriction in its application to criminal justice: Estimator-variable research cannot alter the accuracy of a given witness's account of a real crime; it can only reduce or increase the court's reliance on the witness's testimony. In itself, this may appear to be a small restriction; after all, it provides the court with an empirically derived set of decision rules. However, before that point is addressed, a review of the estimator-variable literature is in order.

Characteristics of the Criminal Event

Crime seriousness. Leippe, Wells, and Ostrom (1978) speculated that a trivial criminal event would produce a lower base rate of accuracy than a criminal event of moderate seriousness. Staging a crime in which a stolen object was believed by subjects to be worth about $50 versus $1.50, Leippe et al. found subjects better able to pick the "criminal" from a six-person photo spread in the case of the high-value object (56% correct identifications) than in the case of the low-value object (19% correct identifications). In addition, Leippe et al. included two cells in which the subject-witnesses did not know the value of the stolen object until 60 sec after the "criminal" had vanished. Under these conditions, crime seriousness

had no reliable effect. This lead Leippe et al. to suggest that crime-seriousness effects are limited to cases wherein the witness(es) knows how serious the crime is *at the time of witnessing*.

Leippe et al. (1978) also suggested that the relationship between crime seriousness and eyewitness accuracy may actually be curvilinear. At some point, a crime may be so serious that it produces a high level of arousal that interferes with information processing or enhances a motivation to not get involved, thereby decreasing eyewitness accuracy. However, Johnson and Scott (Note 3) staged an event in which subject-witnesses overheard a hostile interaction complete with the sound of breaking bottles and crashing chairs. Subsequently, subject-witnesses viewed a criminal-confederate bolt into their room with a bloodied letter opener and blood on his hand and then quickly exit. In another condition, a confederate who was *not* overheard in a hostile interaction entered the subject-witnesses' room with a pen in his hand and grease on his arms and then quickly exited. The general conclusion from their data is that "high arousal *facilitated* the retrieval of information" (Johnson & Scott, Note 3, p. 27). Assuming, as do Johnson and Scott, that subjects believed this horror, it may be beyond the ethical and practical limits of social psychologists to demonstrate the kind of crime severity necessary to adequately test Leippe et al.'s curvilinear hypothesis. Alternatively, perhaps the curvilinear hypothesis is incorrect.

Exposure time. Another characteristic of the criminal event that varies from case to case is the amount of time that the witness is exposed to the relevant stimuli. The research of Loftus (1972) and Hintzman (1976) shows that picture recognition is a monotonically increasing function of functional exposure time. Of course, this monotonic function is not linear, since there are threshold-type effects and ceiling effects. Interestingly, even above threshold, the data appear more like a psychophysical function than a learning curve. This may be related to Schaffer and Schiffrin's (1972) observation that picture recognition does not resemble verbal learning in that providing a rest or rehearsal period following picture exposure does not increase performance.

Complexity. Research by Loftus (1972) and Wells (1972) suggests that the complexity of an event can increase later recognizability. Franken and Davis (1975) corroborated this suggestion with photographs that varied in the complexity of the image. However, as Hastie et al. (Note 2) point out, "While complexity quite plausibly improves recognition (perhaps by increasing the chances that a subject will observe distinctive features), exactly the reverse effect seems to occur in recall situations." For example, Schiffman and Bobko (1974) found that increasing the complexity of an event resulted in greater overestimations of the duration of that event. Temporal overestimation effects have also been reported in staged crimes (e.g., Buckhout, 1974; Marshall, 1969; Johnson & Scott, Note 3).

Familiarity. Familiarity with the physical surroundings that form the context of the criminal event may also affect recall accuracy. Lack of familiarity with the size and/or distance of surrounding objects can produce large distortions in estimates of the size, distance, and acceleration of the perceptual target (Grether & Baker, 1972).

Hastie et al.'s (Note 2) point regarding the distinction between recall and recognition brings up a more general issue. It may be that some factors that enhance facial recognition actually inhibit recognition or recall of other characteristics. For example, Leippe et al. (1978) found no significant effect for the crime-seriousness manipulation on recall for various physical characteristics of the criminal. In fact, the low-seriousness conditions produced nonsignificantly greater recall accuracy than the high-seriousness conditions. It seems only reasonable that direct attention to the facial characteristics of a criminal serves to reduce the amount of time that a witness is attending to the multitude of other characteristics of the event. Consistent with this, Johnson and Scott's (Note 3) data indicate an average within-cell correlation of $-.21$ between their subjects' accuracy of facial identification and errors of commission in recalling the event (errors of omission were not analyzed). This

finding becomes especially interesting to the extent that a cross-examination of a witness results in a discounting of that witness's testimony via showing that the witness gave a fallible account of some lesser characteristic of the event.

Characteristics of the Defendant

Race. Perhaps the most researched characteristic of the defendant is race. Malpass and Kravitz (1969) found that photographs of black faces were more difficult to recognize than those of white faces regardless of the witness's race. Since blacks show less variation in hair color and eye color than their white counterparts, this result may have a basis in stimulus characteristics. However, most cross-racial identification research shows an interaction between race of defendant and race of witness, thereby giving rise to attitudinal and familiarity explanations. In general, white witnesses show a much higher recognition accuracy for white faces than for black faces, whereas black witnesses only show a slight difference or no difference in their ability to recognize whites versus blacks (Cross, Cross, & Daly, 1971; Elliot, Wills, & Goldstein, 1973; Luce, 1974). Adding Orientals as witnesses and defendants increases the complexity of the pattern (see Cross et al., 1971; Luce, 1974), but a general rule that accounts for a large percentage of the data is that within-race identifications are better than cross-racial identifications.

Attractiveness. The social psychological literature shows a number of biases regarding a target person's physical attractiveness (Berscheid & Walster, 1974), and the eyewitness literature is no exception. Cross et al. (1971) reported that recognition memory for facial photographs that subjects perceived as attractive was higher than for the average facial photograph. However, we should note that Cross et al.'s procedure may have helped induce an attractiveness effect. Specifically, Cross et al. told their subjects that they were studying beauty, and the subject's task was to indicate which faces (from a set of photographs) he or she thought were pretty. If the study had been described as being interested in "ugliness" and the subjects had been asked to sort out the *unattractive* pictures, perhaps the results would have been reversed.

Sex. Early work by Howells (1938) suggested that sex of the target person has little effect on identifiability. Cross et al.'s (1971) research, however, suggests that while sex of the target person has no effect on male subjects, female subjects are better at recognizing female faces than they are at recognizing male faces.

Age. Another characteristic of the defendant that might influence accuracy is the defendant's age. Unfortunately, only one study has included the target person's age as a factor (Cross et al., 1971), and for reasons unknown to the current author, the published version of that study did not report a statistical breakdown of that factor.

Characteristics of the Witness

As with characteristics of the defendant, the witness variable that has probably received the most attention is the witness's race. However, because the witness's race generally interacts with the race of the defendant, these results were discussed in the previous section. Sex of the witness was also discussed in the previous section because of its interaction with sex of the defendant.

Perceptual set. The perceptual set of a witness may influence eyewitness accuracy. Borrowing from Craik and Lockhart's (1972) levels-of-processing concept, Bower and Karlin (1974) found that memory for a face is better if, upon initial exposure to the face, individuals are asked to make judgments that require more thought about or analysis of the face (for example, making judgments of honesty as opposed to judgments of sex).

A Critical Evaluation of Estimator-Variable Research

Can we make specific estimates of accuracy? How can estimator-variable research be applied to criminal justice? There are several ways to apply this research, but they entail untested (and sometimes untestable) suppositions. Earlier it was suggested that

one might assess the merits of a given witness by "plugging in" the relevant factors. For example, one may have a checklist such as "What is the victim's race? What is the defendant's race? How attractive was the defendant? What is the witness's sex? How old is the defendant? How severe was the crime? What was the witness's perceptual set? visual context? exposure time?" and so on. But surely any psychologist must realize the futility of such an approach. Can it be assumed that these factors only combine as main effects? Of course not. A defendant's sex would surely not maintain its effect over long exposure times. Using only the research reviewed thus far, one would have to look at a 19th-order interaction! If new research is added showing that the defendant's facial expression affects accuracy (thereby making it a 20th-order interaction), have we learned more and, therefore, can we apply more? Some may answer "yes" to this question, apparently believing it possible to execute such projects in which all levels of the 20 estimator variables (in a few years will it perhaps be 40 variables?) are factorially combined and their interactions and main effects assessed. But, if one only used two levels of each of the 20 variables and created a factorial design to assess all possible interactions there would be 2^{20}, or 1,048,576, cells in the design.

Can we make general statements? Instead of making specific estimates of eyewitness accuracy, perhaps estimator-variable research can yield *general* statements of use to the courts. For example, one might conclude that because the experimental literature shows such poor accuracy rates in general, jurors and judges must be informed of this fact by psychologists. Yet, even this trivial contribution to criminal justice involves risky suppositions. For example, it assumes that judges and jurors currently believe that witnesses are less fallible than they are. Yet *there is no empirical evidence to support the assumption that jurors and judges are overbelieving of witnesses.* Thus, where is the empirical justification for this courtroom intervention? How can we be assured that psychologists' expert testimony will not create jurors and judges who are *less* believing of witnesses than they should be? Perhaps even more problematic is the fact that to make a statement about the fallibility of eyewitness testimony on the basis of estimator-variable research, the expert would supposedly rely on an *average* level of accuracy, a *typical* accuracy rate, or an average accuracy rate of a "typical study." If the expert is able to do so (which is questionable) he or she is also assuming that the literature is unbiased. Yet it can be proposed that the literature is replete with potential biases. For one thing, only one eyewitness study used the concept of "volunteerism" (i.e., only those witnesses who overtly indicate that they could possibly make a positive identification are allowed to see pictures or a lineup). Thus, a study that shows 30% of the witnesses correct and 70% incorrect might find a very high accurate/inaccurate ratio if it selected witnesses as police often do (i.e., testing only witnesses who *freely* indicate that they saw the criminal, freely indicate that they had a sufficient view, and freely volunteer their services).

Another factor that might make the accuracy rates of eyewitness research somewhat unrepresentative is that subject-witnesses almost always know that the "crime" was staged by the time they are given a recall or recognition task. This knowledge could increase the tendency of those who have limited information to go ahead and identify someone. If it were a real crime, however, a false identification would have important implications, and this possibility could reduce the percentage of witnesses who simply guess.

In addition, no research exists in which the subject-witness is also the victim, yet such situations apply to virtually all rape, robbery, and assault cases. It may be that accuracy is extremely high under these conditions, as opposed to the relatively uninvolved, passive-role conditions typical of staged crimes.

There are other reasons why accuracy rates in estimator-variable research may be misleading. The current author suspects that in general, low accuracy rates may be *preferred* among researchers. Specifically, researchers may perceive it as infinitely more interesting, more publishable, and more socially impor-

tant to show low eyewitness accuracy in eyewitness research than to show high accuracy. A high accuracy rate is an implicit null hypothesis that is to be rejected, and the stronger the rejection, the better. Not surprisingly, eyewitness researchers sometimes feel a greater need to dismiss high accuracy, but not low accuracy, as a research fluke. For example, Buckhout (1974), in reporting the results of a staged crime, noted that subject-witnesses were quite accurate in estimating the "criminal's" height and concluded, "This may be because the suspect was of average height" (Buckhout, 1974). Buckhout's observation may be correct, but it introduces the possibility that researchers do not want high accuracy or that high-accuracy data warrant discounting, whereas low-accuracy data speak for themselves.

There is yet another possible reason to question the representativeness of staged "crimes." Personal communications between the author of the current article and other eyewitness researchers suggest that criminal-confederates may be chosen by researchers on an unrepresentative basis. Eyewitness researchers often go to great lengths to insure that their criminal-confederates do not have outstanding features. It is not clear how this concern for "no outstanding features" becomes operationalized in the selection of a criminal-confederate. However, it is possible that the chosen criminal-confederates have physical features that go a long way toward being unrepresentative, and therefore, perhaps criminal-confederates are more difficult to recognize than most criminals.

It may be possible to use estimator-variable research in a way that does not depend upon *direct* generalizations from data. As Schlenker and Bonoma (Note 4) have pointed out, generalizations are made on the basis of corroborated theories, not data alone. Thus, in principle, estimator-variable research has the potential for application to criminal justice through the theoretical development of eyewitness processes. However, in practice, there is little evidence of the kind of theoretical developments that would yield such applicability. Estimator-variable research tells us that the eye is not like a camera, but this concept of perception has been known for centuries, and continued documentation of that fact is probably unnecessary. In the next section, another type of eyewitness research (system-variable research) is described. Instead of simply documenting that the human eye is not like a camera, system-variable research investigates variables that might be useful for *reducing* the discrepancy between the eye and the camera.

A Closer Examination of the System-Variable Literature

As noted earlier, system-variable research differs from estimator-variable research in that the former investigates variables that are manipulable by the criminal justice system, whereas the latter investigates variables whose influence can be estimated (but not controlled) by the criminal justice system. The distinction between estimator- and system-variable research may be important for the eventual establishment of a useful, nonreactionary, applied criminal justice literature. However, before the potential merits and drawbacks of system-variable research are addressed, a review of that literature is in order.

Retention Interval

Time. Apparently, the belief that recall and recognition memory performance decays with the passage of time is so well accepted that no one has bothered to test it with a live, staged crime. Recently, however, Lipton (1977) exposed subjects to a filmed, simulated murder and tested subjects either immediately or after a 1-week delay. As expected, subjects' recall accuracy was poorer (4.3% less) with the delay than when they were immediately tested.

Suggestive interrogation. Eyewitness researchers' primary interest in the retention interval, however, is not time per se, but rather the intervening material. Foremost in this area is the research of Loftus and her colleagues. This research is exemplified by Loftus and Palmer (1974), who showed subjects a film of a traffic accident and varied the format of an intervening interrogation.

When asking subjects about the speed of the cars, Loftus and Palmer either used the question "About how fast were the cars going when they *hit* each other?" or substituted the phrase "smashed into" for "hit." One week later, subjects who were earlier asked the question with the phrase "smashed into" were more likely to indicate that they saw broken glass than were subjects who were earlier asked the question with he verb "hit" in the sentence. Loftus, Altman, and Geballe (1975), and Loftus, Miller, and Burns (1978) have shown similar distortions in recall as a function of other types of intervening interrogation procedures.

Composite drawings. Recognition memory performance may not only be hindered by disguised suggestions on the part of an interrogator but also by an open task that is largely structured by the witness. Hall and Ostrom (Note 5) presented subjects with a facial photograph, and subjects were later required to try to identify the person from a corporal lineup. However, during the time interval between the initial exposure to the face and the subsequent lineup identification, some subjects worked with an artist to create a composite drawing of the to-be-recognized person's face. The results indicated that irrespective of whether the person to be recognized was present in the lineup, the subjects in the composite-drawing conditions made more errors (average = 50% errors) than did subjects in the no-composite-drawing conditions (average = 31% errors).

Mug shots. In addition to composite drawings, it may be that the process of examining mug shots can lead to subsequent misidentifications in a lineup. Brown, Deffenbacher, and Sturgill (1977) exposed subjects to five strangers, with instructions to the subjects that they would later be required to identify the strangers from a lineup. Before the lineup task, however, subjects viewed a set of mug shots that included some of the original strangers and some new faces. In the subsequent lineup test, the percentage of "incidental" members (i.e., faces that were seen only in the mug shots) who were falsely identified was more than double the percentage of false identifications of completely new faces (i.e., seen in neither the original set nor the mug shots). This result suggests that exposure to mug shots themselves may debilitate subsequent recognition memory performance.

Testing

Question structure. Perhaps the most frequently researched of the system variables, testing structure, can have powerful effects on the witnesses' performance. Lipton's (1977) research in which subjects were shown a filmed murder indicated that completely unstructured testimony (i.e., free elaboration without the use of any questioning) produced greater accuracy (91%) than any other type of questioning. Open-ended questions yielded somewhat greater accuracy (83%) than either leading questions (72%) or multiple-choice questions (56%). Each of these question types was significantly different from the others. The quantity of the testimony, however, tended toward the reverse patterning (21%, 32%, 79% and 75% for each of the preceding question types, respectively). Similar results showing that accuracy declines with question structure and specificity have been reported by Borst (cited in Whipple, 1909), Snee and Lush (1941), Marquis, Marshall, and Oskamp (1972), and Marshall (1969). Yet, Lipton's (1977) research also shows that because specific questions can be either negatively biased (i.e., suggesting an incorrect response) or positively biased (i.e., suggesting a correct response), the degree of difference in accuracy between specific and nonspecific questions is attenuated by question bias.

Lineup instructions. Lineups, a multiple-choice type of interrogation, are a particularly appealing focus of research because of the obvious ability of police investigators to control characteristics of a lineup and the pervasive utilization of such evidence in court. Hall and Ostrom (Note 5) found that the instructions delivered to a witness prior to a lineup task can influence false identifications. Specifically, Hall and Ostrom found that telling the witness that the suspect "is in the lineup" led to more false identifications than did telling the witness that the suspect "may or may not" be in the lineup.

Lineup structure. Characteristics of the lineup itself can also influence accuracy. Wells et al. (Note 6) have outlined an empirical technique for helping to ensure the fairness of a lineup. They point out that witnesses are often motivated to pick "someone" out of a lineup because of witnesses' frequent belief that the police have a good set of reasons for organizing the lineup and that the police have amassed evidence against one of the lineup members. If the witness can detect which lineup member is suspected by the police, it may increase the chances of the witness choosing that person. One can imagine an extreme example of such a bias wherein the witness observed a tall, black man commit a crime. If the subsequent lineup is composed of one tall, black man and five others who are either Caucasian, female, or short, the witness can clearly discern whom the police suspect. Although this is an extreme example, Wells et al. point out that any lineup has both a "functional size" (i.e., the number of *feasible* lineup members) and a nominal size (number of persons in lineup). The functional size can be calculated using a simple role-playing paradigm. The role-playing paradigm utilizes "mock witnesses" who are given a general description of the suspect as described by (a majority of) the witness(es). Usually this would include race, height, weight, sex, and age. The mock witnesses are then given a picture of the corporal lineup (which must be provided by the prosecution for the defense counsel) with instructions to "choose the person whom you believe the police suspect." Functional size is then calculated at N/X, where X is the number of mock witnesses choosing the defendant and N is the total number of mock witnesses who make a choice. Thus, a lineup in which half of the mock witnesses choose the defendant has a functional size of two, irrespective of the lineup's nominal size. Similar arguments regarding biases in police lineups have been offered by Doob and Kirshenbaum (1973).

While Wells et al.'s (Note 6) technique for assessing functional size can account for some types of lineup bias, other types of lineup bias do not enter into this analysis. For example, intentional or unintentional cuing by the official(s) in charge of conducting the lineup may influence witness responding. The influence of nonverbal cuing in lineup situations has been demonstrated in a study by Fanselow and Buckhout (Note 7). Fanselow and Buckhout showed subjects a 21-sec, silent, color film of a staged mugging and shooting incident. Subsequently, subjects were individually shown six photographs in which one photograph (target picture) was associated with a "positive treatment" (investigator establishes eye contact with subject-witness, leans forward and smiles), a negative treatment (investigator looks away from subject-witness, frowns, leans backward), or a neutral treatment (no unusual treatment of picture). Each picture was associated with each treatment an equal number of times. After going through all six pictures individually, the subject-witness was requested to go back and look at them again and choose the "mugger." The results revealed that both the positive and negative treatments resulted in more choices of the target picture (23% and 16%, respectively) than did neutral treatment (5%). Note, however, that these percentages suggest that the effect may be primarily due to an inhibition-of-choice effect for the neutral treatment condition in this study, leaving the precise interpretation unclear.

An Evaluation of System-Variable Research

How can system-variable research be applied to the betterment of criminal justice? One possibility is to use it in the courtroom as an estimator variable. For example, an expert could suggest to the court that the conditions of the lineup, courtroom procedures, or length of time between the original event and subsequent testing is so great that the witness is quite likely to be wrong. However, any such courtroom statement, no matter how it is phrased, would imply a likelihood of accuracy with regard to the case at hand. Suppose, for example, that the expert observed a biased lineup. Can he or she say that it is likely that the witness was incorrect because past research shows strong effects for biased lineups? No. If all of the measurable and unmeasurable influencing factors

for that specific case were known (lighting, exposure time, etc.) it may be that the probability of accuracy would be .95 without a biased lineup and .92 with a biased lineup. Biased lineups may have consistently debilitating effects, but likelihood of accuracy depends on too many factors in a given case for any semblance of reasonable estimation. In other words, *this* use of system-variable research entails the same problems of applications as did the estimator-variable research. As discussed earlier, discounting eyewitness testimony is a practice that is based upon questionable derivations from the eyewitness literature (overinduction) and fails to provide an alternative source of evidence.

The alternative use of system-variable research is to manipulate the relevant variables so as to *reduce the inaccuracies of witnesses.* For example, system-variable research can be used to advocate short witness-testing intervals, fairer lineups, reduced use of composite drawings, and so forth. This gives the criminal justice system empirically derived tools with which to better the criminal justice process. It provides an alternative to current practices without advocating the elimination of eyewitness testimony as a rule of evidence. Using system-variable research in this manner can contribute to criminal justice without the practice of post hoc discounting of an eyewitness.

There is one problem that afflicted the application of estimator-variable research that may also apply to the application of system-variable research, namely, the issue regarding statistical interactions. However, this may be less problematic for system-variable applications than it is for estimator-variable applications. In applying estimator variables to a specific case, *all* types of interactions are important. For example, the effect of high versus low crime severity on eyewitness accuracy could be enhanced, moderated, or reversed by another variable. Any of these three types of interaction would be important in determining the likely accuracy of a witness for courtroom presentation. However, in applying system variables, the enhancement and moderation-type interactions present no special problems. For example, the effect of a short versus long incident–test interval may be moderated or enhanced by another variable, but a problem exists only if the directional influence is *reversed* (i.e., a long interval facilitates memory more than a short interval) by the presence of another variable.

Conclusion

It would be bad advice to suggest a halt to estimator-variable research. However, in undertaking an *applied* project, it is incumbent on a researcher to demonstrate the applied utility of an eyewitness study. In this regard, estimator-variable research may not be highly fruitful. This statement does not speak to any specific study reviewed in the estimator-variable section. Some of those studies were not intended to be applied, and some of those studies may have more applied utility than their membership in that category suggests. Yet, the overall conclusion is maintained: System-variable research in eyewitness identification may, as a general rule, have greater applied utility for criminal justice than does estimator-variable research.

On the other hand, there may be certain types of estimator-variable research that can circumvent the application problems that have been listed. For example, instead of focusing on the *situation* as the unit of postdictive analysis, efforts could be made to use the *individual* as a unit of analysis. Specifically, self-report information (e.g., various measures of the witness's confidence in his or her identification) may be diagnostic of accuracy. This cognitive approach assumes that situational factors prior to and during the crime (estimator variables) affect some measurable response (e.g., confidence) in a manner that is statistically related to the accuracy factor. Unfortunately, only three studies have measured any such cognitive storage (see Brown et al., 1977; Leippe et al., 1978; Wells, Lindsay, & Ferguson, in press), and the postdictive utility of the cognitive measure (confidence) has been unproductive.

Reference Notes

1. Buckhout, R. *Nobody likes a smartass: Expert testimony by psychologists.* Paper presented at the annual convention of the American Psycho-

logical Association, Washington, D.C., September 1976.

2. Hastie, R., Loftus, E. F., Penrod, S., & Winkler, J. D. *Eyewitness testimony: Review of the psychological literature.* Unpublished manuscript, Harvard University, 1977.
3. Johnson, C., & Scott, B. *Eyewitness testimony and suspect identification as a function of arousal, sex of witness, and scheduling of interrogation.* Paper presented at the annual convention of the American Psychological Association, Washington, D.C., September 1976.
4. Schlenker, B. R., & Bonoma, T. V. *Fun and games: The validity of games for the study of conflict.* Unpublished manuscript, University of Florida, 1976.
5. Hall, D. F., & Ostrom, T. M. *Accuracy of eyewitness identification after biased or unbiased instructions.* Unpublished manuscript, Ohio State University, 1975.
6. Wells, G. L., et al. *Guidelines for empirically assessing the fairness of a lineup.* Manuscript submitted for publication, 1978.
7. Fanselow, M. S., & Buckhout, R. *Nonverbal cueing as a source of biasing information in eyewitness identification testing.* Unpublished manuscript, Brooklyn College, 1976.

References

Berscheid, E., & Walster, E. Physical attractiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11). New York: Academic Press, 1974.

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

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

Buckhout, R. Eyewitness testimony. *Scientific American,* 1974, *321,* 23–31.

Craik, T. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* 1972, *11,* 671–684.

Cross, J. F., Cross, J., & Daly, J. Sex, race, age and beauty as factors in recognition of faces. *Perception & Psychophysics,* 1971, *10,* 393–396.

Devlin, Rt. Hon. Lord Patrick. *Report to the Secretary of State for the House Department on Evidence of Identification in Criminal Cases.* House of Commons, April 26, 1976. London: Her Majesty's Stationery Office.

Doob, A. N., & Kirshenbaum, H. M. Bias in police lineups—Partial remembering. *Journal of Police Science and Administration,* 1973, *1,* 287–293.

Elliot, E. S., Wills, E. J., & Goldstein, A. G. The effects of discrimination training on the recognition of white and oriental faces. *Bulletin of the Psychonomic Society,* 1973, *2,* 71–73.

Franken, R. E., & Davis, J. Predicting memory for faces from rankings of interestingness, pleasingness, complexity, figure-ground and clarity. *Perceptual and Motor Skills,* 1975, *41,* 243–247.

Grether, W. F., & Baker, C. A. Visual presentations of information. In H. P. van Cott & R. G. Kinkade (Eds.). *Human engineering guide to equipment design.* Washington, D.C.: U.S. Government Printing Office, 1972.

Hintzman, D. L. Repetition and memory. In G. L. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10). New York: Academic Press, 1976.

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

Leippe, M., Wells, G. L., & Ostrom, T. Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology,* 1978, *3,* 345–351.

Levine, F. J., & Tapp, J. The psychology of criminal identification: The gap from Wade to Kirby. *University of Pennsylvania Law Review,* 1973, *121,* 1079–1131.

Lipton, J. P. On the psychology of eyewitness testimony. *Journal of Applied Psychology,* 1977, *62,* 90–95.

Loftus, E. F., Altman, D., & Geballe, R. Effects of questioning upon a witness's later recollections. *Journal of Police Science and Administration,* 1975, *3,* 1–2; 165.

Loftus, E. F., Miller, D. G., & Burns, H. J. Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, *4,* 19–31.

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

Loftus, G. R. Eye fixations and recognition memory. *Cognitive Psychology,* 1972, *3,* 525–557.

Luce, T. S. Blacks, whites and yellows: They all look alike to me. *Psychology Today,* November 1974, 105–108.

Malpass, R. S., & Kravitz, J. Recognition for faces of own and other race. *Journal of Personality and Social Psychology,* 1969, *13,* 330–334.

Marshall, J. *Law and psychology in conflict.* New York: Doubleday-Anchor, 1969.

Marquis, K. H., Marshall, J., & Oskamp, S. Testimony validity as a function of question form, atmosphere and item difficulty. *Journal of Applied Psychology,* 1972, *56,* 167–186.

Schaffer, W. O., & Schriffrin, R. M. Rehearsal and storage of visual information. *Journal of Experimental Psychology,* 1972, *92,* 292–296.

Schiffman, H. R., & Bobko, D. J. Effects of stimulus complexity on the perception of brief tem-

poral intervals. *Journal of Experimental Psychology*, 1974, *103*, 156–159.

Shepard, R. N. Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 156–163.

Snee, T. S., & Lush, D. E. Interaction of narrative and interrogatory methods of obtaining testimony. *Journal of Psychology*, 1941, *11*, 229–236.

Wells, G. L., Lindsay, R. C. L., & Ferguson, T. J. Accuracy, confidence, and juror perception in eyewitness identification. *Journal of Applied Psychology*, in press.

Wells, J. E. Encoding and memory for verbal and pictorial stimuli. *Quarterly Journal of Experimental Psychology*, 1972, *24*, 242–252.

Whipple, G. M. The observer as reporter: A survey of the "psychology of testimony." *Psychological Bulletin*, 1909, *6*, 153–170.

Whipple, G. M. The obtaining of information: Psychology of observation and report. *Psychological Bulletin*, 1918, *15*, 217–248.

Woocher, F. D. Did your eyes deceive you? Expert psychological testimony on the unreliability of eyewitness identification. *Stanford Law Review*, 1977, *29*, 969–1030.

Received March 24, 1978 ■