Does Knowledge that a Crime Was Staged Affect Eyewitness Performance?¹

DONNA M. MURRAY AND GARY L. WELLS²

University of Alberta

The staged crime has generally been considered a valid paradigm for assessing eyewitness identification accuracy. Yet the typical procedure informs the witness before the lineup task that the event was staged. The current study manipulated whether or not witnesses were informed that the witnessed crime was staged (information variable). Two other variables, the presence or absence of the perpetrator in the photo-lineup and the physical similarity of lineup members, were manipulated to assess any possible interactions with the information variable. A theft was staged individually for 184 unsuspecting witnesses who were then randomly assigned to be either informed that the theft was staged or not informed. Next, all witnesses were met by a uniformed security officer who showed them a photo-lineup of six individuals. Half of the witnesses viewed a perpetrator-present lineup with either high or low physical similarity between the foils and the perpetrator and half viewed a perpetrator-absent lineup with either high or low physical similarity. Three confidence measures, one cooperation measure, and an Embedded Figures Test score were obtained. The results showed that (a) the information variable did not affect the likelihood that a witness would attempt an identification; (b) the information variable did not have a main effect on the particular choices made by the witnesses nor did it interact with lineup similarity, but it did interact with the presence-absence variable; and (c) the information variable did not affect witness confidence either as a main effect or interaction effect, but was relatively consistent in affecting the correlation between confidence and accuracy such that the correlation was lower for witnesses who were not informed. The interaction between the information variable and the presence-absence variable on accuracy was due to the uninformed witnesses being less accurate than the informed witnesses in the perpetrator-present condition. The results suggest that

¹This research was conducted as a masters thesis by the first author under the direction of the second author. Thanks are due to Grant Grobman for his role as the policeman, Richard Blackwell for his excellent portrayal as the perpetrator, Ric Reid for his professional advice on acting and makeup, C. M. Goode of Campus Security for providing authentic equipment, and Alinda Friedman for her advice on use of factorial analyses.

²Requests for reprints should be sent to Gary L. Wells, Department of Psychology, University of Alberta, Edmonton, Alberta, T6G 2E9.
accuracy among actual eyewitnesses may be lower than obtained in the typical staged crime procedure and that accuracy-confidence correlations may be overestimated by the typical staged crime.

The use of staged events for understanding the psychology of eyewitness testimony dates back to the early part of the century (e.g., Munsterberg, 1908). The use of unsuspecting eyewitnesses to a live event has special advantages over the more traditional memory experiments when it comes to matters of generalizing to eyewitness evidence. The traditional experiment on memory for faces, for example, uses still photographs as the original stimulus and directs the subject to take a particular set for encoding. A staged event (e.g., staged crime), however, preserves most of the ecological factors of importance by using dynamic, live events and letting the witnesses’ encoding of information below follow its own natural incidental course.

How well does the staged crime represent its real-world counterpart in terms of eyewitness identification? Because we can’t make veridical assessments of real-world eyewitness accuracy, we must rely in part on the face validity of the staged crime paradigm. In this regard there is one characteristic of the typical staged crime procedure that may undermine its ecological representativeness. Specifically, in the typical staged crime procedure the eyewitness is informed prior to the identification task that the crime was simulated. There are good reasons for informing the eyewitness at this point. Practically speaking, it is extremely difficult to maintain the deception for a lineup identifications task. The witness expects a number of things prior to an identification attempt, such as questioning from police, time for apprehension of a suspect, and appearance at a police department.

Nevertheless, the knowledge that the crime was staged prior to the identification attempt may be quite important. The informed witness now knows that her/his identification has no serious consequences for herself/himself or for the accused. It is thus a sheer test of memory without the possible interfering effects of emotional arousal, there is no fear of reprisal from identifying a criminal, no fear of sending an innocent person to jail, no court appearance, etc. Both anxiety level at the time of retrieval and the consequences of reporting a retrieval are different for the real-world witnesses than they are for the standard witnesses of staged crimes. Thus, we might expect differences in accuracy for informed vs. uninformed eyewitnesses and/or a criterion shift (e.g., less willingness to attempt an identification).

The use of informed eyewitnesses is the general rule in staged crime research (e.g., Buckhout, Figueroa, & Hoff, 1975; Leippe, Wells, & Ostrom, 1978; Lindsay, Wells, & Rumpel, 1981; Malpass & Devine, 1981; Wells, Lindsay, & Ferguson, 1979). There are, however, two exceptions. Malpass and Devine (1980) staged a vandalism and kept eyewitnesses blind to the deception through
the subsequent identification task. However, there were no informed witnesses with which to compare the performance of uninformed witnesses. A study by Sanders and Warnick (in press), though, was designed to allow such a comparison. They found no difference between witnesses who knew prior to the identification that the act (cheating) was staged vs. those who were uninformed of this fact. The Sanders and Warnick study, therefore, lends credibility to the previous research that used informed witnesses.

There are two reasons for being hesitant to accept the null effect conclusion from the Sanders and Warnick (in press) study. First, the cost of identifying the cheater was relatively trivial for the identified person. Identifying the cheater resulted in the cheater being eliminated from competition, a competition that someone who needed to cheat is unlikely to win. Cost to the witness also was trivial since there was no indication that the witness would have to get involved (e.g., go to court and testify). Second, there were no “background” variables manipulated in the study that might show the information variable interacting with some other variable. Because the issue is one of paradigm validation, we ultimately want to know if the informed-witness paradigm is valid for examining the effects of manipulated variables. If the effect of a variable is negated or reversed by the knowledge that the crime was staged, then, for that variable, the standard eyewitness paradigm would not be valid.

The background variables chosen for this experiment were the presence-absence of the perpetrator in the photo lineup (i.e., presence-absence variable) and the similarity of lineup foils to the accused. Presence-absence was selected because in the real world any lineup can be either a criminal present or a criminal absent lineup. Any statement about variables affecting eyewitness performance must take into account these two possible states of the world. The similarity variable was included because of its controllable role as a system variable (see Wells, 1978) and the fact that previous theoretical and empirical treatments have accorded it considerable importance (e.g., Doob & Kirshenbaum, 1973; Lindsay & Wells, 1980; Wells, Leippe, & Ostrom, 1979). It is generally believed that increasing the physical similarity of the lineup foils (nonsuspects) to the suspect helps protect the innocent suspect (Doob & Kirshenbaum, 1973; Wells, Leippe, & Ostrom, 1979). Recent work by Lindsay and Wells (1980) suggests more precisely that increases in similarity reduce identifications of both the innocent suspect and the actual offender, with the former being reduced more than the latter. The result is an overall increase in the diagnostic value of lineup identifications (Wells & Lindsay, 1980).

Given these concerns, we designed a study to test some general hypotheses. First, informed and uninformed subjects may differ in regard to their willingness to make a choice from the lineup. Second, the information variable may interact with some other variable such as presence-absence and/or the similarity variable. Third, the information variable may affect the confidence
of eyewitnesses and/or the relationship between confidence and accuracy.

In addition, we wanted to test various ideas regarding the confidence of eyewitnesses and, in particular, the relationship between confidence and accuracy. This is an important issue because there is evidence that the relationship between confidence and accuracy may be weak or nonexistent (e.g., Brown, Deffenbacher, & Sturgill, 1977; Leippe et al., 1978; Lindsay et al., 1981; Wells, Lindsay, & Ferguson, 1979; Wells, Ferguson, & Lindsay, 1981). This weak relationship between eyewitness confidence and eyewitness accuracy has been shown to mislead triers of fact as they rely heavily on eyewitness confidence to infer eyewitness accuracy (Wells et al., 1981). It is possible, however, that other measures of confidence may be better suited to the task of predicting accuracy. For example, a measure taken prior to the witnesses’ choice might be more highly correlated with accuracy than the more common measure taken after the identification (Leippe, 1980). It is also possible that several measures in combination may predict accuracy better than any one measure in isolation. Therefore, several measures of confidence were included in this study. In addition, some empirical evidence (Goodenough, 1976) suggests that a personality measure (Embedded Figures Test) may be related to memory in social situations. For this reason, the Embedded Figures Test was included as another possible predictor of accuracy.

Method

Subjects

One hundred and eighty-four introductory psychology students participated (one at a time) in the experiment as part of their course credit. In addition, 10 subjects were eliminated because they did not believe the cover story; 12 subjects were eliminated because they knew one of the confederates or a member of the lineup; and 1 subject was eliminated because of experimental foul-up.

Design

Half of all subjects viewed a lineup that contained the confederate perpetrator and the other half viewed a lineup in which the perpetrator was replaced by the innocent suspect. Half of the subjects in these conditions were informed that the theft was staged, while the other half were led to believe that the theft was real. Finally, half of the subjects in these conditions viewed a lineup in which the nonsuspects’ physical appearances were high in similarity to the perpetrator and half viewed a lineup in which the nonsuspects were low in physical similarity to the perpetrator. Thus the design was a 2 (perpetrator present vs. absent) X 2 (witnesses informed or uninformed that the theft had been staged) X 2 (high vs. low similarity lineup) between-subjects factorial.
Procedure

The experiment required an experimenter, a uniformed confederate policeman and a confederate perpetrator dressed as a maintenance worker. Although the experimenter and confederates interacted with the subjects, they were blind to the subjects' experimental group membership. The subjects were told that they would be playing a TV game and taking an Embedded Figures Test (Witkin, 1971) as part of an investigation of game strategies. A modified version of the Embedded Figures Test was used (i.e., Section I was skipped, and subjects were given 5 minutes to complete as many problems as possible from Sections II and III).

The subject was told the room number for the experiment and told that he/she would be allowed to practice in private for a few minutes. As the subject entered the room, the perpetrator was in the final process of disconnecting the TV game from the TV. The perpetrator appeared startled and hurriedly put the TV game under his jacket. The perpetrator then dropped his screwdriver, picked it up while looking at the subject, and brushed past the subject as he ran out of the room. As the perpetrator ran down the hall, he looked over his shoulder three times to see if he was being followed. After a few minutes, the experimenter arrived and discovered that the TV game was missing. The experimenter told the subject that a theft had occurred and that Campus Security would be called.

After supposedly calling Campus Security, the experimenter told the subject-witness that "Campus Security says other things have been stolen from this building recently, and Campus Security thinks there may be a connection between the thefts and all the maintenance work that is being done in the building lately. They say that they have pictures of the maintenance men who have been working in the building and have been waiting for an opportunity to show the pictures to someone. Campus Security says that they'll be right over and that they want you to wait for them." (Maintenance workers were actually working in the building at the time.) It was suggested to the subject-witness that she/he might as well take the Embedded Figure Test while waiting for Campus Security.

Just before leaving the subject, the experimenter turned the page of the Embedded Figures Test booklet so that the book was open to the first figure, which the subject was required to find. The presence or absence of a small pencil mark on this page determined whether the witness was in the uninformed or informed condition. Informed subjects were told, "There was no real theft. The theft you saw was staged by us. We'll explain later, but for now we want you to continue to behave the same way you would have behaved if we hadn't told you that the theft was staged. The Campus Security person will be arriving soon, and he's also one of us." Uninformed witnesses were not, of course, given such information.
Five minutes later, the Embedded Figure Test was stopped by the arrival of the confederate policeman who was dressed in a Campus Security uniform. The "policeman" carried an envelope containing six pictures. Upon arrival, the policeman asked each subject, "Do you think you can make an identification?" The subject's answer was noted in the policeman's notebook. This response was used as a predecision measure of witness confidence. The policeman then told the subject, "It's possible that none of these pictures is the man you saw. In fact, the thief may have been impersonating a maintenance man so that he could move in and out of rooms in the building without being noticed." The policeman then handed the subject the envelope containing the six pictures. The same ordering of the pictures was used for all subjects.

The pictures had been placed in the envelope previously so as to produce one of the four possible lineups (i.e., perpetrator present, high similarity; perpetrator present, low similarity; perpetrator absent, high similarity; perpetrator absent, low similarity). This means that half of all subjects viewed a lineup that contained the confederate perpetrator (i.e., perpetrator present lineup); and the other half viewed an identical lineup except that the perpetrator was replaced by the innocent suspect (i.e., perpetrator absent lineup). The innocent suspect was selected on the basis of similarity ratings in pilot data. The perpetrator and the innocent suspect never appeared together in a lineup. Half of the subjects viewed a high similarity lineup and half viewed a low similarity lineup. High and low similarity was also determined on the basis of pilot work. The envelopes, which contained the lineups, were arranged in a random order, and the policeman took the top envelope to the subject.

After the subject made his/her decision, the policeman asked either, "How sure are you that this is the man?" or "How sure are you that none of these is the man?" depending on the subject's response. The response to this question was jotted down in the policeman's notebook and was used as a postdecision measure of confidence. The policeman then asked, "Would you be willing to come down to the station and view a live lineup?" That response was also recorded and was used as a behavioroid measure of the witness's willingness to cooperate. In addition, if an identification was made, the policeman asked, "Would you be willing to sign a statement to the effect that this was the man that you saw in the room?" This response was used as a behavioroid measure of confidence. [Unfortunately, both behavioroid measures were introduced after 36% of the subjects had already participated in the study; and therefore, responses for these measures are not available for all subjects.] The experimenter then returned and the policeman left.

The experimenter debriefed the subjects. At the start of the debriefing, the experimenter asked the uninformed subjects broad, general questions. The questions gradually became more narrow and more focused on the possibility that some form of deception had occurred. This method of debriefing
allowed the experimenter to ascertain whether the subject believed the cover story and also allowed the subject to discover the deception for himself. The experimenter's debriefing was followed by a debriefing from the perpetrator. In the second debriefing, the confederate perpetrator gave the subject further details about the research and obtained a promise of secrecy from the subject.

Rating scales were constructed for the confidence and cooperation measures. Neither the scale constructor nor the raters had knowledge of the subjects' experimental conditions. Agreement between the two raters, independently conducted, exceeded 95%. Discrepancies were resolved through discussion.

Results

Analysis of Subjects' Identification Choices

A 2 (presence-absence) × 2 (information) × 2 (similarity) × 3 (choice) $\chi^2$ was used to analyze the subjects' choices. The three possible choices were selection of the suspect, selection of a foil, or a refusal to make an identification.

Results of this analysis revealed a significant present-absent × choice interaction ($\chi^2(2) = 15.19, p = .001$), a significant similarity × choice interaction ($\chi^2(2) = 6.16, p < .05$), and a significant present-absent × information × choice interaction ($\chi^2(2) = 6.48, p < .05$). No other significant effects were observed with this analysis.

The proportions yielding the present-absent × information × choice interaction are shown in Table 1. A comparison of each proportion for informed witnesses with its counterpart for uninformed witnesses revealed one significant difference and one marginally significant difference. Specifically, when perpetrator present lineups were used, uninformed witnesses were significantly more likely to choose a foil than were informed witnesses ($Z = 2.66, p < .01$) and

<table>
<thead>
<tr>
<th></th>
<th>Criminal present</th>
<th></th>
<th>Criminal absent</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Choice of</td>
<td>Choice of</td>
<td>No choice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>suspect</td>
<td>foil</td>
<td></td>
</tr>
<tr>
<td>Informed witnesses</td>
<td>.43</td>
<td>.13</td>
<td>.43</td>
<td>.11</td>
</tr>
<tr>
<td>Uninformed witnesses</td>
<td>.26</td>
<td>.37</td>
<td>.37</td>
<td>.11</td>
</tr>
</tbody>
</table>
marginally less likely to choose the perpetrator \( Z = 1.75, p = .08 \). Informed and uninformed witnesses did not differ in choices of suspect or choices of foils in the perpetrator absent conditions. Informed and uninformed witnesses also did not differ in willingness to make an identification in either the perpetrator present or perpetrator absent conditions.

The present-absent \( X \) choice interaction appears to the fact that the suspect was chosen more frequently in the perpetrator present conditions (i.e., when the suspect was the perpetrator) than in the perpetrator absent conditions \( Z = 3.86, p < .001 \). In addition, witnesses more frequently made no choice when the perpetrator was absent than when the perpetrator was present \( Z = 2.51, p < .02 \). There was no difference in choice of foils as a function of the present-absent variable. Thus, the pattern of the present-absent \( X \) choice interaction indicates that subject witnesses were operating above chance in discriminating the presence-absence of the perpetrator in the lineup.

The similarity \( X \) choice interaction was due to a significantly higher choice of foils in the high similarity conditions than in the low similarity conditions \( Z = 2.47, p < .02 \). The similarity variable did not, however, significantly affect refusals to identify or choices of the suspect \( Z's = 1.33 \) and 1.05, respectively, \( p's > .18 \).

**Predictors of Accuracy**

**Confidence.** The first three confidence measures (predecision confidence, postdecision confidence, and willingness to view a live lineup) were subjected to a 2 (present-absent) \( X \) 2 (information) \( X \) 2 (similarity) \( X \) 2 (identification vs. no identification) weighted means analysis of variance. The latter variable, i.e., whether the witness identified someone or not, is a subject variable. It was included because of recent speculations that there may be main effect differences between choosers and nonchoosers among eyewitnesses in confidence (e.g., Leippe, 1980). Because the fourth confidence measure (willingness to sign a statement) was taken only of choosers, it was subjects to a 2 (present-absent) \( X \) 2 (information) \( X \) 2 (similarity) ANOVA. Results of these analyses revealed only two significant effects. There was a main effect for witnesses to express more postdecision confidence in the perpetrator present than in perpetrator absent conditions, \( F(1,167) = 4.136, p = .044 \), and witnesses expressed greater willingness to sign a statement in the perpetrator present than in perpetrator absent conditions, \( F(1,39) = 10.25, p = .003 \).

It was believed that the information variable might affect the confidence-accuracy correlation, even though it had no effect on the mean level of confidence. Therefore, we calculated the confidence-accuracy correlation for informed witnesses separately from uninformed witnesses. This was done for each of the four confidence measures. These correlations are presented in
Table 2. Table 2 reveals no evidence of a confidence-accuracy correlation for the uninformed witnesses. However, for informed witnesses two of the confidence measures were significantly correlated with accuracy, namely the traditional postdecision confidence measure and the measure of willingness to sign a statement.

It was also hypothesized that a predecision measure of confidence might be better correlated with accuracy than postdecision measures. This hypothesis was based on the viewpoint that self-perception processes emanating from choice behavior (a la Bem, 1965) serve to lower the confidence-accuracy correlation (see Leippe, 1980). There is no support for this hypothesis in Table 2 as predecision confidence was clearly unrelated to accuracy. This observation, along with the fact that choosers and nonchoosers were equally confident in their decisions, portends poorly for the self-perception viewpoint.

Thus, there is no evidence that confidence measures taken prior to identification are better than postdecision measures. There is evidence, however, suggesting that confidence may be better correlated to accuracy for informed rather than uninformed witnesses.

Embedded figures. The Embedded Figures Test scores were not significantly correlated with accuracy of witnesses’ choices ($r = .091, p = .266$). This correlation remained nonsignificant when an analysis was done for chooser only (i.e., only the witnesses who attempted an identification ($r = .151, p = .192$)).
Multiple regression. A multiple regression using all predictors of accuracy (i.e., the four confidence measures plus the Embedded Figures Test scores) was calculated to examine the maximum predictive power achieved through use of these measures. The resultant correlation was +.543 ($p < .05$). The multiple regression, however, was greater when calculated for informed groups only ($r = .837$, $p < .05$) than when calculated for uninformed groups ($r = .405$, n.s.) and the difference between these correlations was significant ($Z = 2.06$, $p < .05$). Thus, as with the individual correlations on the confidence measures, there is evidence that the predictors of accuracy for informed witnesses work more poorly for uninformed witnesses.

Discussion

The current data suggest that there may be significant differences in the results obtained from the traditional staged crime paradigm, wherein witnesses are informed prior to the lineup that the event was staged, and actual eyewitness settings where such briefings do not occur. Specifically, for perpetrator present lineups we found differences in the likelihood of choosing foils and of identifying the suspect. In addition, the data suggest that eyewitnesses’ confidence in their lineup decisions, once hoped to be a convenient correlate of accuracy, may be even less useful in actual eyewitness settings than the typical informed-witness procedure led us to believe.

If our information variable captures the ecological differences between real-world witnesses (who are not informed that a crime was an experiment) and the typical staged-crime experiment (in which witnesses are so informed), then certain conclusions may be warranted.

First, factors such as confidence or personality measures may better predict eyewitness accuracy in a traditional staged-crime study than they can in real-world setting. Our data show that accuracy-confidence correlations, both individually and in multivariate combination, are lower for witnesses who have not been informed that the event was staged than for witnesses who are informed of such. As discussion before, the traditional staged-crime paradigm tends to show very weak correlations between confidence and accuracy. However, the more realistic, uninformed witness paradigm seems to produce even lower correlations, correlations that clearly qualify as being useless in any practical sense. It seems clear that any success with the traditional (informed) staged-crime paradigm in predicting accuracy with the use of subject variables might eventually need to be corroborated with an uninformed witness procedure. The informed witness procedure may paint a picture that is somewhat too rosy for the real-world eyewitness situation.

It is important to note that the similarity variable affected choices of foils and that this effect did not interact with the information variable. This suggests
that the effects of similarity might be safely explored with an informed witness procedure. In other words, we obtained no evidence suggesting that the effects of similarity would be different for informed vs. uninformed procedures.

Because the information variable interacted with the presence-absence variable, however, some caution should be exercised in generalizing from the informed procedure. The data indicate that uninformed witnesses were marginally less likely to choose the perpetrator and significantly more likely to choose a foil in the perpetrator-present lineups. Thus, there should be some value placed on the general plan of corroborating results obtained from the informed witness procedure by using an uninformed witness procedure.

We are not willing at this time to suggest that all research use the uninformed witness procedure. The uninformed procedure is an extremely difficult procedure to implement. The maintenance of deception throughout an identification task involves difficult logistical efforts. It is stressful for experimenters and the confederate policeman who must carry out acting roles above and beyond the traditional domain of experimentation. There may also be ethical issues involved in this form of deception, especially as regards the obvious stress experienced by a witness who believes that his or her identification has profound implications for a suspect. The traditional informed-witness procedure, in contrast, involves mild, indirect deception. The witness in the informed procedure is never explicitly told that a crime occurred, and learns almost immediately after the witnessed event that the event was staged.

We suggest that the informed witness procedure remain the principal method for exploring eyewitness variables. However, once a variable has been found to be robust and reliable with the informed procedure there should be an attempt to validate it further with an uninformed procedure. In this way we can use a relatively safe and effective paradigm to weed out the most promising variables. Of course, there is a risk in our suggestion in that the informed witness procedure might find some variable ineffective, yet that variable would have had robust effects with the uninformed procedure.

REFERENCES


