0-315-03657-5

National Library of Canada — Bibliothèque nationale du Canada
Canadian Theses Division — Division des thèses canadiennes
Ottawa, Canada
K1A 0N4

51552

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

Donna Marie Murray

Date of Birth — Date de naissance

6-6-46

Country of Birth — Lieu de naissance

U.S.A.

Permanent Address — Résidence fixe

6122 BeaverDam Way NE
Calgary, Alberta T2K 3W7

Title of Thesis — Titre de la thèse

Effects of Knowledge that a crime was staged on Eyewitness Performance

University — Université

Univ. of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

M.A.

Year this degree conferred — Année d'obtention de ce grade

1981 - Spring

Name of Supervisor — Nom du directeur de thèse

Gary Wells

Date

1-16-81

Signature

Donna Marie Murray

NL-91 (4/77)

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE

THE UNIVERSITY OF ALBERTA

Effects of Knowledge that a Crime was Staged on Eyewitness Performance

by

© Donna Marie Murray

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF Master of Arts

Psychology

EDMONTON, ALBERTA

SPRING 1981

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Donna Marie Murray

TITLE OF THESIS Effects of Knowledge that a Grime was Staged on Eyewitness Performance

DEGREE FOR WHICH THESIS WAS PRESENTED Master of Arts

YEAR THIS DEGREE GRANTED December 16, 1980

(SIGNED) Donna Marie Murray

PERMANENT ADDRESS:

6122 Beaverdam Way N.E.

Calgary, Alberta, T2K 3W7

DATED : December 16, 1980

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Effects of Knowledge that a Crime was Staged on Eyewitness Performance submitted by Donna Marie Murray in partial fulfilment of the requirements for the degree of Master of Arts.

Supervisor

Date December 16, 1980

Abstract

One hundred and eighty-four subjects observed a theft, made an identification from a lineup, and answered other questions about the event. Half the subjects were informed that the theft was staged prior to viewing a lineup, while the other half believed that the theft was real throughout the experiment (i.e., informed vs. uninformed eyewitnesses, respectively). The physical similarity of the lineup foils to the criminal (i.e., lineup similarity) was also manipulated as well as the presence or absence of the criminal in the lineup. The subjects' selection from the lineup was recorded. In addition, a predecision measure of confidence, a postdecision measure of confidence, a behavioroid measure of cooperation, a behavioroid measure of confidence, and an Embedded Figures Test score were obtained. When the criminal was in the lineup, uninformed witnesses were less accurate than informed witnesses. The information variable did not affect witness accuracy when the criminal was not in the lineup. For choosers only, the postdecision confidence measure and the behavioroid confidence measure were significantly correlated with accuracy. Multiple regression was used to predict the informed group's accuracy from the three measures of confidence, the measure of cooperation, and the Embedded Figures Test. A similar multiple regression was calculated for the uninformed group. The multiple regression for the informed group was significant, and it was also significantly higher than the regression for the uninformed group. The multiple regression for the uninformed group was not significant. Lineup similarity had a marginal effect on accuracy. Lineup similarity did not interact with the information variable. There was no difference between informed and uninformed subjects on willingness to make an identification.

Acknowledgements

I would like to thank my confederates, Richard Blackwell and Grant Grobman for their work; and I would like to thank C. M. Goode of Campus Security for providing authentic equipment. I would also like to thank Ric Reid for professional advice on acting and on the effective use of theatrical makeup. I would like to thank Eugene Couch for his help with computer programs for the data analysis. I would also like to thank Alinda Friedman for her many suggestions concerning the thesis, especially those regarding Chi Square.

Most importantly, I would like to thank my advisor, Gary Wells for many stimulating conversations, for broadening my knowledge of experimental design, for improving the quality of the research, and for his overall supervision, which made the thesis work an enjoyable experience.

Table of Contents

Chapter | Page

List of Tables

List of Appendixes

It has long been recognized that psychology might be usefully applied to eyewitness testimony (Munsterberg, 1908). Yarmey (1979) notes that both Freud, in 1906, and Watson, in 1913, discussed the importance of psychology for eyewitness testimony and jury decision making. This issue has been explored through both facial recognition research and eyewitness research.

In facial recognition research, the subject first views photographs of individuals. A second set of photographs is then introduced. This second set contains pictures of individuals who were in the original set of photographs and pictures of individuals who were not in the original set of photographs. The facial recognition task is to view the second set of photographs and identify those individuals who were also in the original set of pictures.

In eyewitness research, a crime is staged and then individuals are asked to view a lineup (or a set of pictures). The subjects are expected to either identify the criminal or state that the criminal is not in the lineup. As will be discussed below, there are problems in generalizing to the real world criminal justice system from either facial recognition research or from eyewitness research.

A. Facial Recognition Research

Recently, investigators have attempted to examine the variables that affect facial recognition. This research has obvious implications for the criminal justice system especially for understanding the variables that operate in eyewitness identifications. Unfortunately, comparison between experiments is difficult because of stimulus differences between experiments. For example, some experimenters use schematic faces, others use photos, and still others use faces made by selecting from many possible eyes, noses, etc. Despite difficulties generated by stimulus variability, some findings are generally agreed upon.

The ability of subjects to verbally describe a particular face is not correlated with the ability to recognize the face when it is presented again (Goldstein, Johnson, & Chance, 1979). In general, when experimenters use Caucasian faces and Caucasian subjects and immediate recall, recognition accuracy rates vary between 54% and 96% with an average of about 79% (Clifford and Bull, 1978, p. 93).

In addition to the generally high accuracy rates, investigators have found that only-minor changes in accuracy occur when delays of up to one week are interposed between the first and second viewing of the face. For example, Laugherty, Fessler, & Lenorovitz (1974) found that similar accuracy was obtained whether a four minute or a one week delay was used. Brown, Deffenbacher, & Sturgill (1977) found that, even with a two day interval between original viewing and testing for recognition, subjects were able to correctly identify 96% of the faces. However, Brown et al.'s subjects performed only slightly better than chance when it came to recalling where they had seen the face.

It is also agreed that the upper portion of the face, particularly the forehead, hairline, and eyes, is more important for recognition than the lower portion of the face (Yarmey, 1979, p. 116). However, it is also agreed that faces are not recognized solely on the analysis of individual features (Yarmey, 1979, p. 116).

Other factors that influence facial recognition are judged attractiveness, distinctiveness, and likeableness of the face (Light, Stuart, & Hollander, 1979; Yarmey, 1975; Yarmey, 1978). Cross, Cross, and Daly (1971) found that faces independently rated as attractive are more likely to be identified; while Shepherd & Ellis (1973) report that faces rated "high" on attractiveness and faces rated "low" on attractiveness were more easily identified after five weeks than faces rated "medium" on attractiveness.

Research with facial recognition has also established that some faces are more difficult to discriminate than others. Cohen & Carr (1975) and Light et al. (1979) found that, if attention is not called to the dimension of distinctiveness during the initial viewing of the face, faces rated "low" in distinctiveness are more difficult to recognize; and when they are used as decoys, they result in more misidentifications. Goldstein, Stephenson, and Chance (1977) note that some faces are never mistaken for the target face while other faces are often mistaken for the target.

Laugherty et al. (1974) concluded that the critical variables determining whether a correct identification will occur are the number of faces from which the selection is to be made and the homogeneity of the faces. In one of Laugherty et al.'s (1974) experiments, homogeneity was based on the amount of overlap on nine physical characteristics – hair color, hair length, age, build, eye color, glasses, moustache, beard, and length and shape of sideburns. In another experiment, homogeneity was based on similarity ratings obtained from the subjects. Both homogeneity and number of faces are inversely related to the probability of a correct identification.

It is also known that recognition is substantially altered by simply changing non-permanent aspects of the individual's appearance. For instance, removing a beard caused accuracy to drop from 97.5% to 42% (Laugherty & Fowler, 1977). Laugherty & Fowler have also analyzed misidentifications. They found that when a decoy is mistakenly identified for the target, it is usually the case that the decoy has long hair and/or a beard and/or a moustache. They argue that subjects rely heavily on changeable aspects of facial appearance. Expression is also easily changed and both Galper (1970) and Sorce & Campos (1974) have discovered that if the target's expression is not the same on both photo presentations, accuracy decreases.

Bower & Karlin (1974), Warrington & Ackroyd (1975), and Strnad & Mueller (1977) all found that asking subjects to make a judgment as to whether a face appeared pleasant or honest resulted in significantly better recognition than asking the subjects to make a more superficial judgment regarding the face (e.g., size or sex). Yarmey (1978) found that when the subjects were required to make attributions regarding the face (e.g., happy), a two-second viewing of the photo was sufficient to insure 89% correct recognitions. In addition, Yarmey (1979) states that memory is positively correlated with the number of different dimensions on which these judgments are made

It has been argued that attention to these "deeper" dimensions, rather than the dimension per se, is the important factor in the improved recognition (Bower & Karlin, 1974; Yarmey, 1975). This view is supported by Winograd (1976) who demonstrated that making a judgment with respect to the dimensions of honesty and pleasantness improved performance recognition regardless of whether the subject decides that the target possesses the quality or decides that the target does not possess the quality. All the above has obvious implications for any lineup. For example, deeper dimensions at the time of the crime should increase the witness's accuracy when he/she views the lineup.

The improved accuracy in facial recognition that results from attention to one or more relevant dimensions also occurs in the more general paradigm of discrimination learning. Gibson (1967 p. 99) argues that discrimination learning is facilitated when properties on which the stimuli differ are emphasized in training. This suggests the possibility that requiring subjects to make judgments regarding types of noses would increase accuracy by the same means as the judgments regarding pleasantness, etc. (i.e., by calling attention to a relevant dimension). However, research indicates that this is not the case. Patterson & Baddeley (1977) found that judgments regarding personality characteristics of the target resulted in better recognition performance than judgments about facial features (e.g., type of nose). The effect of these judgment tasks on memory are generally included under a conceptual framework of levels of processing originally developed by Craik and Lockhart (1972). The underlying assumption of the levels of processing approach is that a stimulus is processed through a series of stages with different types of information being extracted from the stimulus at successive stages. The more levels or stages through which the stimulus is processed the better its retrievability from memory. Presumably, tasks requiring judgments such as honesty require these deeper levels of processing more than do tasks requiring more superficial judgments such as sex. Loftus (1979), on the other hand, suggests that judgments regarding personality characteristics may improve recognition relative to judgments regarding facial features because the former requires an analysis of the entire face as opposed to only a single feature of the face. The former is also more "semantic". That is the subject must relate the face to more abstract qualities, which may result in a more fully elaborated memory representation.

B. Eyewitness Literature and Comparisons Between the Eyewitness Literature and the Facial Recognition Literature

There are similarities between results obtained in the facial recognition experiments and results obtained in eyewitness experiments. For example, Brown et al., Study 2, (1977) found that when subjects view a crime, then view mug shots, and then view the lineup, the subjects confused the confederate criminal with faces seen only in the mug shots. This confusion is analogous to Brown et al.'s Study 1 (1977) finding that subjects could recognize a previously seen face, but had trouble in determining where they had seen the face.

In addition, Lindsay & Wells (1980) found that increasing the physical similarity of the lineup members to a witness's verbal description of the criminal (i.e., increasing similarity) decreased selections of the criminal. This is comparable to Laugherty et al.'s (1974) finding that increasing homogeneity between faces decreases recognition accuracy. However, facial recognition research usually defines accuracy differently than does eyewitness research. In facial recognition experiments, accuracy is simply the number of correct identifications divided by the total number of choices; but that portion of the eyewitness research that deals with the problem of similarity between lineup members uses a more ecologically–valid definition of accuracy.

In the real world, it is generally the case that all members of the lineup except the suspect are known to be innocent. Since no criminal proceedings will ensue from the selection of someone other than the suspect, accuracy can be calculated as the number of correct identifications in a lineup that includes the actual criminal divided by the number of identifications of the person in the criminal–absent lineup who is designated as the innocent suspect (Lindsay & Wells, 1980). Although the resultant ratio of correct identifications (when the target is present in the choice set) to incorrect identifications (when the target is absent from the choice set) leads to a somewhat different definition of accuracy than is traditionally used, the definition is one that more closely approximates the operational characteristics of the criminal justice system.

When Lindsay & Wells (1980) used this definition of accuracy, they were able to show that increases in similarity between lineup members (i.e., what

Wells, Leippe, & Ostrom (1979) refer to as increasing functional size) resulted in a decline in accurate identifications (i.e., identifications of the criminal), but an even greater decline in false identifications (i.e., identifications of the innocent suspect). Thus, choices of the correct face could be said to decrease as similarity among the choice set increases (as suggested by the facial-recognition literature), but the similarity variable has an even stronger effect on reducing choices of the wrong face (i.e., innocent suspect) when the correct face is absent from the choice set.

Lindsay & Wells (1980) argue that when their definition of accuracy is used, high lineup similarity (functional size) increases accuracy by necessitating a deeper level of analysis and necessitating more elaborate use of memorial processes rather than surface analysis. The term "functional size" is meant to characterize the number of functional or *plausible* members in the lineup, based on physical similarity, and is distinquished from nominal size or the mere *number* of persons in the lineup. The functional size of a lineup is calculated by providing individuals who have not seen the criminal (i.e., mock witnesses) with a description of the criminal and asking the mock witnesses to select the criminal from the lineup on the basis of the description. The total number of identifications are divided by the number of selections of the suspect. Thus, a lineup in which 50% of the mock witnesses choose the suspect has a functional size of two, regardless of its nominal size.

C. Problems with Generalization from Facial Recognition Research to the Real World

As previously noted, Lindsay and Wells (1980) point out that in the criminal justice system, lineups are usually composed of only one suspect. All other members of the lineup are known to be innocent. This means a false identification is only significant in the case where an innocent suspect is identified as the criminal. In terms of laboratory research, the comparable situation would be a lineup in which the confederate criminal is absent; and in addition, the selection of any individual in the lineup other than the previously designated innocent suspect is not counted as a false identification since no prosecution would result from this selection. Lindsay and Wells (1980) defined the innocent suspect as the member of the lineup who the witnesses most often misidentified as the criminal. Lindsay & Wells (1980) used this method of calculating false identifications and manipulated functional size (i.e., physical similarity between the members of the lineup). When the functional size of the lineup was 2.44 and the criminal was not in the lineup, Lindsay & Wells found that 31% of those who attempted an identification choose the innocent suspect. However, when the criminal absent lineup with a functional size of 1.1 was used, 70% of those who attempted an identification, identified the innocent suspect.

Accuracy rates obtained using the eyewitness procedure tend to be lower than the accuracy rates obtained in some facial recognition tasks (e.g., Deffenbacher et al., 1975). Wells, Lindsay, & Ferguson (1979) found that 26% of the identifications made by witnesses were false identifications (i.e. someone other than the criminal confederate was identified as the criminal) while Leippe et al. (1978) found that more than 50% of the identifications made by witnesses were false identifications. Neither of these studies used a criminal absent lineup (i.e., It was always possible to choose the criminal.)

Of course, accuracy rates are not necessarily comparable across eyewitness experiments because accuracy is affected by a multitude of variables (e.g., viewing time and frequency, salience, expectations, etc.; Loftus, 1979). Accuracy could be affected by a change in any of these variables. For example, Lindsay, Wells, & Rumpel (1980) were able to affect accuracy rates by manipulating trivial situational variables. They employed three groups. In the high-accuracy condition, the subject interacted with the criminal for 20 seconds and the criminal did not wear a hat. In the moderate-accuracy condition, the subject interacted with the criminal for 12 seconds and the criminal wore a hat; but the hat did not cover the criminal's hair. The low-accuracy condition was identical to the moderate condition except that the hat was pulled down so that the criminal's hair and ears were covered. Of the subjects who made identifications in Lindsay et al.'s high, moderate, and low situational accuracy conditions, 26%, 50%, and 67% respectively were false identifications. In addition, witnesses in the low situational accuracy condition were not less likely to attempt an identification or less likely to be willing to testify in "court" than witnesses in the high situational accuracy condition.

Despite low accuracy rates, subjects tend to be highly confident of their identifications. In fact, accuracy and confidence have been found to be either uncorrelated (Brown et al., 1977; Leippe et al., 1978) or only weakly correlated (Lindsay et al., 1980; Wells et al., 1979).

Investigation has also shown that the magnitude of the witness's confidence in his/her identification does not reflect situational variables designed to raise or lower identification accuracy (Lindsay et al., 1980). Similarly, although functional size has been shown to affect accuracy, a witness's confidence in his/her accuracy is not altered by changes in the functional size of the lineup (Lindsay & Wells, 1980).

Poor accuracy combined with high confidence does not portend well for the use of eyewitness testimony in legal proceedings. This is especially true in view of research that indicates that jurors' belief of a witness is heavily determined by the witness's confidence (Lindsay et al., 1980; Wells et al., 1979; Wells, Lindsay, & Tousignant, 1980).

Despite similarities between facial recognition research and eyewitness identifications that occur in the real world and between facial recognition research and eyewitness research, there are several reasons why it seems unlikely that the facial recognition procedure is adequate for generalization to the recognition task required of eyewitnesses in the real world.

First, arousal is known to affect performance. At high levels of arousal, performance declines (Yerkes-Dodson Law). This is expecially true of complex tasks (e.g., person identification), which involve the integration of several thought processes. In addition, as arousal increases, attention may shift from the general environment to only a subset of the environment (e.g., a shift from criminal and weapon to only the weapon; Easterbrook, 1959). Since facial recognition subjects do not believe that they have witnessed a crime, it is likely that they may be less aroused during input than individuals who have witnessed a crime and, therefore, encode the situation differently.

Second, the facial recognition subject is often told that he/she will be asked to identify the target person(s) later. This means that he/she can actively process the target. But, in everyday experience, observations are often not planned and deliberate so that the witness is often not in a position to intentionally study and store information regarding the criminal. This "incidential" nature of criminal observation can occur either because the witness does not realize that a crime has occurred until the criminal has gone or because the witness is so surprised by the crime that strategic forms of observation do not occur.

Experiments that ask the subject to study the face and determine whether a particular quality (e.g., distinctiveness) is present also differ from the real world in that witnesses are not given such instructions. In addition, it's argued that recognition memory is not necessarily improved by the mere presence of a particular quality. Instead it's the process of deciding whether the quality exists that improves recognition memory. Therefore, this aspect of facial recognition research can not possibly be applicable to crimes that do not provide adequate opportunity or motivation for analyzing the criminal's appearance.

Nevertheless, it's possible that, under certain conditions, real world witnesses may also make dispositional judgments about the criminal. For example, it may be automatic, to judge the criminal's meanness, desperation, etc. Unless specifically required to do so, it's unlikely that the subjects in facial recognition research regularly make dispositional judgments about the target. Facial recognition research that specifically requires subjects to make such judgments may more closely approximate the real world than other types of facial recognition research. This rationale is what was used by Leippe, Wells, & Ostrom (1978) in the crime seriousness study. Leippe et al. argued that a more serious crime produced deeper processing of the criminal's intentions ("Is he/she really a thief?"), which resulted in better recognition performance.

The facial recognition experiments also differ from the real world in that often the identical stimulus (e.g., photo) is used for both the original viewing and the later recognition test. In addition, when a different picture is used for the recall task, research indicates that minor changes such as expression affect the subject's accuracy. It's highly unlikely in the lineup that one could duplicate everything about the criminal, including his expression during the crime. In the real world, many aspects of the criminal have changed prior to the second identification. This makes it risky to generalize to the real world from facial recognition tasks in which the same photograph is used for the original viewing and the later recognition task.

A final major difference between facial recognition research and the real world is that subjects in facial recognition experiments are not exposed to live targets. A live person is much more dynamic than a picture. It is possible that encoding of a dynamically changing individual is more difficult than the more limited recognition tasks required in the facial recognition experiments.

The eyewitness procedure in which a subject actually believes that he/she has witnessed a crime does not suffer from any of the above problems. Therefore, it seems more appropriate for investigating the problems of real world eyewitness identifications.

D. Possible Problems in Eyewitness Research

It's possible that the currently used eyewitness laboratory procedure, like the facial recognition task, is not a good model of the actual eyewitness situation. In the laboratory, the witness is told that the crime was staged prior to making an identification. This creates a major distinction between the laboratory witness and the witness to an actual crime. Standard eyewitness research assumes that the important factor is the subject's belief at the time of encoding that a crime occurred. After the initial encoding, it is assumed that no differences result from the subject's awareness that the crime was simulated.

However, this assumption is not empirically grounded and seems counter-intuitive. In addition, this position is in contrast to considerable evidence that memory is affected by conditions at the time of retrieval. Abernathy (1940) demonstrated that exam performance decreased when the room used for testing was different than the room used for teaching. Lipton (1977) found that asking specific questions, instead of letting the subject tell what happened in his/her own words, increased quantity of recall but reduced accuracy. Loftus & Zanni (1975) showed that question wording can affect recall. Asking subjects if they saw "the" broken head light is more likely to produce a yes response than is asking if the subject saw "a" broken head light. In a similar vein, Loftus & Palmer (1974) were able to change estimates of speed by changing the verb in the sentence "About how fast were the cars going when they "contacted" each other"? Subjects estimates of speed were affected by whether the verb was "contacted", "hit", "bumped", "collided", or "smashed".

Possible differences in accuracy

Both anxiety level at the time of retrieval and the consequences of reporting a retrieval are different for the real world witness than they are for the laboratory witness. It's possible that this may make generalizations to the real world questionable.

In the real world, witnesses are under pressure both to identify the criminal and to be correct in their identifications. This type of pressure, especially the pressure created by the need to be correct, can be expected to cause more anxiety than is experienced in the laboratory where making a decision, and more importantly, making a correct decision, is not so critical.

The case for a performance decrement due to increased anxiety at the time of retrieval of straightforward. Performance on complex tasks, such as eyewitness identifications, is known to be seriously hampered by high levels of arousal. This relationship between performance and arousal was previously discussed with respect to information input differences between facial recognition subjects and eyewitness subjects.

The consequences of making an identification are clearly not the same for both the laboratory subject and the real world witness. The laboratory witness knows that no serious consequences for either himself or the suspect will result from his/her identification, while the real witness knows that very serious consequences may result from his identification.

The difference in consequences can be seen by comparing Table II with Table I. In the laboratory, the negative consequences of making an identification are minimal and are identical to the negative consequences of not making an identification. This is not a caution-inducing situation and, therefore, a witness may feel free to take his/her best guess as to the identity of the criminal.

In the real situation, any identification is costly since one must then appear in court. Because the consequences in the real world are more serious, one may hypothesize that a real witness will be both more cautious and more aroused than a laboratory witness in making any identification. In addition, it seems likely that many real witnesses would be particularly cautious to avoid Box 3 (i.e. sending an innocent person to jail). This means a stricter criterion for identifying a suspect may be used in the real world situation.

Although a strong rational case can be made for a change in performance due to a change in consequences, the empirical case for the consequences variable is less strong than the case for the arousal variable. A major source of inferential support for the differential consequences hypotheses comes from signal detection theory (Egan, 1975).

Because of methodological differences, signal detection theory is not directly applicable to the standard eyewitness research. However, signal detection theory is a type of decision theory and, therefore, in some ways, analogous to the eyewitness paradigm. Since the two tasks are similar, it seems reasonable, as a first approximation, to use signal detection theory metaphorically to determine areas of importance for research in the eyewitness paradigm.

Signal detection theory breaks decisions into two components - the subject's ability to detect a signal (e.g., the criminal) in a signal plus noise situation (e.g., the lineup) and the subject's willingness to say that he/she has detected the signal. Being willing to say that a signal was present can be broken down into the subject's expectations regarding whether the signal is present and the consequences of reporting the signal.

In terms of eyewitness research, the subject's expectations include variables such as the suggestion that the suspect is in the lineup or differential treatment of a particular lineup member (e.g., using a color picture for the suspect and black and white pictures for the foils). Expectation variables are known to be effective in influencing whether a particular eyewitness identification will occur (Fanselow & Buckhout, 1976; Malpass & Devine, 1980 thus providing confidence in the metaphorical use of signal detection theory and, in particular, providing confidence that the consequences of an identification may influence eyewitness behavior.

The possibility that identification consequences may be an important factor in eyewitness identifications is not only suggested by similarities between signal detection theory and the eyewitness paradigm. Malpass and Devine (1980) found that witnesses were more likely to make an identification (not necessarily an accurate one) if they believed that the punishment for the crime would be severe rather than trivial. Malpass and Devine (1980) believe that when the crime is viewed as a serious infraction, witnesses make more identifications because the inconvenience that an identification would cause the witness is offset by the importance of identifying the criminal. This is similar to saying that the consequences of making an identification influence whether an identification is made.

If witnesses actually are using a different criterion in the real world than in the laboratory, then it's possible that the accuracy rates obtained in the laboratory are not generalizable to the real world. If the witness employs a stricter criterion, he/she will need to be more sure before risking an identification. With a strict criterion, both the criminal and innocent suspect would be identified less often. However, one might expect that the innocent suspect, more often than the criminal, would not meet the stricter criterion for identification. Since accuracy is based on the ratio of selections of the criminal to selections of the innocent suspect, accuracy would increase under the above circumstances.

Possible differences in confidence ratings and in confidence-accuracy relationship

A stricter criterion also raises the possibility that there may be differences between the real world and the laboratory with respect to witness confidence. The lower the witness's confidence in the identification, the more likely it is that the identification is one that would not meet a stricter criterion. Therefore, a stricter criterion may remove low confidence identifications in the real world. This means a real world witness who makes an identification may be more confident than a laboratory witness who makes an identification. In addition, it's possible that the real world environment might operate to raise or lower confidence independently of any changes in accuracy (Leippe, 1980). Since confidence and accuracy may both behave differently in the real world than they do in the laboratory, it's possible that the laboratory confidence-accuracy relationships (or lack thereof) do not apply to the real world.

Possible differences between predecision and postdecision confidence

Since laboratory witnesses have typically been very confident of their decisions, it might be argued that the difference in the consequences of making an identification are not that important. However, the confidence measure is often collected after an identification has been made. Bem (1972) argues that people use their own behavior to make inferences about their beliefs. This suggests that once an identification has been made an individual may infer that he/she must have been confident because he/she made an identification. It may be that predecision confidence measures are better related to accuracy than are postdecision confidence measures.

Possible differences in willingness to make an identification

It's also possible that willingness to make an identification, independent of any changes in accuracy or confidence, may be different in the real world than in the laboratory. Real world witnesses may believe that they could make an identification but be unwilling to do so. Conversely, witnesses in the real world may be operating under pressures, which do not exist in the laboratory (e.g., feelings of social responsibility). These pressures may affect the witnesses' willingness to make an identification so that the real world witness may feel compelled to identify someone especially as crime seriousness increases. The Malpass and Devine (1980) study provides some support for this hypothesis by demonstrating that willingness to make an identification is not a constant, but rather increases with consequences to the offender. However, none of Malpass and Devine's subjects were aware that they were participating in an experiment; and therefore, the Malpass and Devine study can not provide information about possible differences between the laboratory witness and the real world witness.

Because of the possible differences between identifications in the laboratory and in the real world, it is important to experimentally determine whether the standard experimental eyewitness procedure is applicable to the criminal justice system.

Possible interaction between laboratory variables and real world variables

The problems of generalizability extend further than just questions of whether accuracy, confidence, and willingness to identify a suspect are different in the two situations. It's possible that variables present in the real world may interact with standard eyewitness variables such as functional size. For example, it may be that the functional size effect results from the willingness of laboratory witnesses to identify the lineup member who most closely fits their remembrance of the criminal. When functional size is low the innocent suspect is the best fit, but when functional size is high, the foils are also good fits and draw choices away from the innocent suspect. However, in the real world, it may be that witnesses are not willing to simply select the closest fit. They may only be willing to make an identification if there is an absolute fit rather than a relative fit. In the real world, increasing functional size may have no effect on decreasing identifications of the innocent suspect. It may be that independent variables that have been manipulated in the laboratory (e.g., functional size) would result in totally different patterns of behavior in the real world. This would mean that the laboratory results are not only different from the real world results, but also that they are not proportional to results obtainable in the real world. If the results are different in magnitude, but comparable in direction, then it would only be necessary to scale the laboratory results in order to apply them to the real world. However, if the results are different and, in addition, not comparable, generalization from the laboratory to the real world would not be justified.

One way of empirically determining whether the laboratory is an adequate representation of the real world is to compare subjects who participate in a standard eyewitness experiment with subjects who participate in an identical experiment except that the latter are not informed that the crime was staged until after the identification is made (i.e., manipulation of the information variable).

In addition, if any differences should develop, it would be necessary to determine whether the laboratory results could be scaled so that they adequately represent the real world or whether the information variable interacts with other variables so that the performance of the typical laboratory witness is unrelated to the real world. The possibility that the information variable interacts with standard eyewitness variables can be tested by the inclusion of variables already known to affect eyewitness accuracy.

Both presence-absence of the criminal and functional size are known to affect eyewitness accuracy. The inclusion of functional size is particularly advantageous since research (Lindsay & Wells, 1980) has shown that accuracy increases when the functional size is increased from 1.10 to 2.44. However, it has not yet been determined whether the relationship between functional size and accuracy is linear. The inclusion of functional size would provide evidence as to whether even higher functional size lineups produce further increases in accuracy.

If no significant differences occur between the informed and uninformed groups, then the standard laboratory procedure is supported as a means of generalizing to the real world. Of course, it is still possible that the two situations differ with respect to variables other than the presence-absence variable and the functional size variable (e.g., crime seriousness). If a significant difference occurs between the informed and uninformed groups and an interaction does not occur between the information variable and one of the other variables, the standard laboratory procedure is still validated although the results of laboratory studies will have to be appropriately adjusted before they can be applied to the real world.

Certain types of interactions would indicate a serious problem with generalization. If, for example, the information variable serves to either eliminate or reverse the effect of the functional size variable, then future work on functional size should be done with uninformed witnesses. If, however, the information variable moderates or accentuates the functional size effect, then future work on functional size may be valid within an informed-witness paradigm as long as appropriate precautions of interpretation are made (e.g., that the effect might be less/more robust with uninformed witnesses).

Summary

A considerable amount of research has been done on the factors that affect facial recognition. However, it is doubtful whether facial recognition research is generalizable to the task of the eyewitness in the real world. Eyewitness laboratories have also generated a considerable quantity of research, but the fact that subjects become aware that the crime was staged prior to viewing the lineup raises the possibility that eyewitness research is also not generalizable to the task of the real world eyewitness. It is possible to determine whether the standard laboratory eyewitness research procedure generalizes to the real world by experimentally manipulating whether or not the subject is informed that the crime was staged and, at the same time, manipulating a second variable known to affect eyewitness accuracy.

Seven separate questions are being tested in this experiment. First, accuracy rates may be different for subjects who believe that they are participating in an experiment (i.e., informed group) than they are for subjects who do not possess this information (i.e., uninformed group). Second, confidence may not be the same for the informed and uninformed groups. Third, the relationship between confidence and accuracy may be different for the informed and uninformed groups. Fourth, willingness to make an identification may differ for the two groups. Fifth, predecision confidence ratings may be more closely related to accuracy than are the standard postdecision confidence ratings. Sixth, functional size has an inadequate empirical base at the higher levels of functional size. This experiment attempts to determine whether higher levels of functional size replicate Lindsay & Wells (1980). Seventh, functional size may not have the same effect for informed and uninformed groups.

I. Method

A. Subjects

One hundred and eighty-four introductory psychology students participated in the experiment as part of their course credit. In addition, ten 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 one subject was eliminated because of experimental foul-up.

B. Design

A 2 (present-absent) x 2 (informed-uninformed) x 2 (functional size) x 3 (Choice of Subject) design was used. The present-absent variable was necessary because a one-suspect lineup means that correct identification can only be made from a criminal present lineup while false identifications can only occur with a criminal absent lineup. (See Lindsay & Wells, 1980.) The functional size values for the four lineups (i.e., criminal present, high functional size; criminal present, low functional size; criminal absent, high functional size; and criminal absent, low functional size) were 3.14, 1.57, 22.0, and 3.67 respectively. These values were determined by pilot work in which a general description of the criminal was compiled. This description and a set of six pictures were given to 88 individuals (22 per lineup). None of these 88 individuals had seen the criminal confederate. The individuals were asked to select the picture that most closely resembled the description. Functional size is the total number of identifications per lineup (i.e., 22) divided by the number of identifications of the suspect. (See Appendix A for the description used to determine functional size).

The innocent suspect was the picture that received the highest similarity rating to the criminal in the pilot study. The high functional size lineup was composed of those pictures that had received the second through sixth highest similarity ratings. The low functional size lineup was composed of those pictures that had received the lowest scores for similarity to the criminal. The ten pictures comprising the middle of the similarity ratings were discarded.

The criminal appeared in 50% of the lineups. In the other 50% of the lineups the criminal was absent and the innocent suspect took the criminal's place in the lineup. The criminal and innocent suspect were never present in the same lineup.

Once the lineups had been formed, the functional size determinations were confirmed by further pilot work in which additional similarity data was collected. Similarity ratings were obtained for all possible pairwise combinations of the suspects and foils of each lineup. A repeated measures ANOVA was calculated for the subset of similarity ratings in which a suspect is compared to a foil. The ANOVA contained a main effect for presence-absence ($F(1, 84) = 11.949$, $p = .001$) and a main effect for size ($F(1, 84) = 7.614$, $p = .007$). There was no interaction between the presence-absence variable and the size variable ($F(1, 84) = .070$, $p = .791$). The main effect for size indicates that the pictures represent two distinct levels of functional size (i.e., high and low). The main effect for presence-absence indicates that the functional sizes of the criminal absent lineups are higher than the functional sizes of the criminal present lineups. These main effects are in agreement with the functional sizes previously determined by using mock witnesses. However, the absence of an interaction in the similarity data suggests that the difference between the criminal absent, high functional size lineup and the criminal absent, low functional size lineup may not be as extreme as suggested by the functional size determination.

C. Procedure

The experiment required an experimenter, a confederate policeman, and a confederate thief dressed as a maintenance worker. Although the experimenter and confederates interacted with the subjects, they were blind to the subject's experimental group membership.

The experimenter met the subject and walked with him/her to the corridor, which led to the room where the theft occurred. While walking, the experimenter explained to the subject that the purpose of the experiment was to examine game strategies. The subject was told that he was required to play a TV game and then take an Embedded Figures Test (Witkin, 1971). Based on a review of the literature, Goodenough (1976) has argued that field independence is related to intentional learning in a social environment while field dependence is related to incidental learning in a social environment. It was thought that this eyewitness situation might be an instance of intentional learning in a social environment. A modified version of the Embedded Figures Test was used (i.e., Section I was skipped, and subjects were given five minutes to complete as many problems as possible from Sections II and III). When the experimenter and subject reached the corridor to the room in which the theft was to occur, the subject was told the number of the room and told, "I'm going to give you a few minutes to get warmed-up. You'll probably feel more comfortable practicing without me watching you so I'll check on the subjects taking the Embedded Figures Test and be back in a few minutes".

As the subject entered the room, the thief was in the final process of disconnecting the TV game from the TV. The thief appeared startled and hurridly put the TV game under his jacket. The thief then dropped his screwdriver, picked it up while looking at the subject, and then brushed passed the subject as he ran out of the room. As the thief ran down the hall, he looked over his shoulder three times to see if he was being followed. (Actually, he was looking back in order to give the subject more time to view his face).

Shortly after the theft, the experimenter returned and discovered that the TV game had been stolen. The experimenter became upset and questioned the subject regarding the missing TV game. The questioning contained the implication that the subject may have taken the game. Once the subject supplied the experimenter with the fact that the subject had seen someone take the game, the subject was told that Campus Security would have to be called and that Campus Security would want to speak to the student. It was then suggested that the subject might as well participate in the second part of the experiment while he/she waited for Campus Security.

After supposedly calling Campus Security, the experimenter told the subject, "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". (Mainentance workers were actually working in the building at the time). If the subject had not already done so, he usually mentioned at this point that the thief looked like an electrician.

The subject was then given the Embedded Figures Test. The Embedded Figures Test served four functions. It kept the subject from thinking about the cover story and, thereby, becoming suspicious; it kept the subject busy so he could not socialize in the halls; by preventing rehearsal, it created an effect similar to the passage of time; and it was thought that the Embedded Figures Test might be useful as a predictor of eyewitness accuracy. The experimenter went over the Embedded Figures Test instructions with the subject.

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. If there was a small pencil mark in the upper right hand corner of this page, the subject was assigned to the uninformed condition. If there was no pencil mark, the experimenter told the subject, "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."

Five minutes later, the Embedded Figures 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. The pictures had been previously placed in the envelope so as to produce one of the four possible lineups (i.e., criminal present, high functional size; criminal present, low functional size; criminal absent, high functional size; criminal absent, low functional size). The envelopes were arranged in a random order, and the policeman took the top envelope to the subject.

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 the 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. After handing the pictures to the subject, the policeman turned his back on the subject and faced the wall. Supposedly, this was done to give the witness some privacy while he examined the pictures. Actually, it was done to insure that the policeman remained blind to condition by insuring that he did not see the lineup that was being used.

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 these questions was jotted down in the policeman's notebook and was used as the postdecision measure of confidence. The policeman then asked, "Would you be willing to come down to the station and view a live lineup"? This response was again 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 late additions to the research; and therefore, responses are not available for all subjects. The experimenter then returned, and the policeman left.

The experimenter debriefed the subjects. A tunnelling process of debriefing was used for the uninformed group to allow the experimenter to ascertain whether the subject bought the cover story and also to allow the subject to discover the deception for himself and, thereby, save face if he/she so desired. The experimenter's debriefing was followed by a second debriefing. In the second debriefing, the confederate thief gave the subject further details about the research and obtained a promise of secrecy from the subject.

The responses to the four confidence questions were examined in order to construct a rating scale for the verbal responses. The obtained rating scale is presented in Appendix B. Neither the scale constructor nor the raters had knowledge of the subjects' experimental conditions during these two tasks. Subjects' responses were assigned numbers as indicated in Appendix B. A second, independent rating was obtained for any response that was not specifically stated under the numerical values in the Appendix B scale. For that subset of subject responses that were not specifically listed in Appendix B, the interrater reliability (i.e., perfect agreement between the two judges) for the predecision confidence measure, postdecision confidence measure, willingness to view a live lineup, and willingness to sign a statement was 76%, 86%, 83%, and 100% respectively. In addition, when differences between raters occurred, they were small (i.e., 1 point).

II. Results

For each of the experimental conditions, the proportion of each possible response (i.e., identify suspect, refuse to make an identification, identify foil #1, identify foil #2, etc.) was calculated. These proportions are shown in Table III.

A. Effect of the Information Variable on Accuracy

A 2 (presence-absence) x 2 (information) x 2 (functional size) x 3 (choice) Chi Square was calculated. Catagory membership in the presence-absence, information, and functional size catagories was determined by the experimental condition to which a subject was assigned. Catagory membership for the choice variable was determined by the subject's response to the lineup. Three responses were possible (i.e., selection of either the guilty or innocent suspect, selection of a foil, or a refusal to make an identification). See Table IV for Chi Square Summary Table.

The information x choice interaction from this Chi Square was not significant (Chi Square (2) = 2.494, p = .287). This indicates that the informed group did not differ from the uninformed group in the liklihood of identifying the suspects, the foils, or refusing to make an identification. However, both identifications of the suspect and refusals to make an identification can be either accurate or inaccurate depending on whether a criminal present or a criminal absent lineup was used. In order to determine the effect of the information variable on witness accuracy, it is necessary to examine the presence-absence x information x choice interaction. This interaction was significant (Chi Square (2) = 6.475, p = .039). See Table V. When the criminal was present, the uninformed group selected the suspect (i.e., the criminal) marginally less often than did the informed group (Z = 1.752, p = .08). In addition, when the criminal was present, the uninformed group selected significantly more foils than did the informed group (Z = 2.658, p = .007). The difference in refusals to make an identification was not significant (Z = .662, p = .509).

A separate Chi Square was calculated for the information x choice interaction for the criminal present groups. The interaction was significant (Chi Square (2) = 7.504, p = .023). This indicates that when the criminal is present, the information variable affects the witnesses' choices; and the previously calculated Z scores indicates the direction of the effect (i.e., a decrease in identifications of the suspect and increase in identifications of the foils for the uninformed group relative to the informed group). This means that when the criminal is present, manipulation of the information variable affects witness accuracy.

A separate Chi Square was also calculated for the information x choice interaction for the criminal absent groups. This Chi Square was not significant (Chi Square (2) = .868, p = .648). In addition, when the criminal was absent there were no significant differences between informed and uninformed groups with respect to identifications of the suspect (i.e., innocent suspect), identifications of the foils, or refusals to make an identification (Zs = 0, .907, & .847 respectively; ps = 1.0, .363, & .395 respectively). When the criminal was absent from the lineup, accuracy was not affected by the information variable.

In summary, when the criminal was in the lineup, the uninformed subjects were less accurate than the informed subjects (i.e., identified foils rather than the criminal). However, when the criminal was absent from the lineup, the information variable had no effect on witness accuracy (i.e., no statistical differences between informed and uninformed groups with respect to choices of the suspect, choices of the foils, or refusals to make an identification).

B. Effect of Presence-Absence Variable on Accuracy

There were no predictions made regarding a main effect for the presence–absence variable on accuracy. However, a basic assumption in eyewitness research is that identifications of the suspect will be greater under the criminal present condition than under the criminal absent condition and that refusals to make an identification will be greater under the criminal absent condition than under the criminal present condition.

The overall Chi Square analysis revealed a significant presence–absence x choice interaction (Chi Square (2) = 15.190, p = .0005). Further analysis of this interaction demonstrated that significantly more identifications of the suspect occurred when the criminal was present than when the criminal was absent ($Z = 3.861$, $p = .000$). Also significantly fewer refusals to make an identification occurred when the criminal was present as compared to when the criminal was absent ($Z = 2.510$, $p = .012$). The criminal present and criminal absent lineups did not differ in the number of times a foil was selected ($Z = .818$, $p = .412$).

In summary, the presence-absence x choice interaction indicates that witnesses were significantly more likely to identify the criminal than the innocent suspect. In addition, if the criminal was not in the lineup, witnesses were significantly more likely to refuse to make an identification than they were when the criminal was in the lineup.

C. Effect of Information Variable on Confidence Ratings

It was thought that there might be a qualitative difference between those individuals who actually made an identification and those who did not. On this basis, the subjects were split into choosers (i.e., individuals who identified either the suspect or a foil) and nonchoosers (i.e., individuals who refused to make an identification). A 2 (presence-absence) x 2 (information) x 2 (functional size) x 2 (choosing vs. nonchoosing) weighted means ANOVA was calculated for the predecision confidence measure (i.e., "Do you think you can make an identification")? Similar ANOVAs were calculated for the postdecision confidence measure (i.e., "How sure are you that this is the man?"), and for the behavioroid measure of cooperation (i.e., willingness to view a live lineup). (The behavioroid measure of confidence (i.e., willingness to sign a statement) was only collected from choosers). There were no main effects for choice and no interactions between choice and any other variable. The ANOVAs were collapsed over choice. A 2 (presence-absence) x 2 (information) x 2 (functional size) ANOVA was calculated for the predecision confidence measure, the postdecision confidence measure, the cooperation measure, and the behavioroid measure of confidence. None of these ANOVAs contained either a main effect for information or an interaction between the information variable and the other variables in the experiment.

Some of the collapsed confidence ANOVAs included main effects for the presence–absence variable, thereby, indicating confidence was affected by whether or not the criminal was present. The postdecision confidence ANOVA suggested that witnesses were marginally (F (1, 175) = 2.925, p = .089) less confident when criminal absent lineups were used than when criminal present lineups were used. Willingness to sign a statement was also affected by the presence–absence variable. When a criminal absent lineup was used, subjects were significantly (F (1, 39) = 10.254, p = .003) less willing to sign a statement.

As would be expected, there was no presence–absence effect in the predecision confidence ANOVA (F (1, 176) = .035, p = .851). There was also no presence–absence effect on willingness to view a live lineup (F (1, 110) = 2.656, p = .106).

In summary, there was no evidence that manipulation of the information variable affected the amount of confidence experienced by the witness. However, the presence–absence variable did affect the postdecision confidence measure and the behavioroid confidence measure.

D. Effect of Information Variable on Predictors of Accuracy

There is support for the hypothesis that the confidence–accuracy relationship is affected by the information variable.

Ideally each lineup should contain only one suspect. Based on this ideal, the most ecologically valid way to calculate the confidence–accuracy correlations would be to base the correlations on only those individuals who identify the suspect because only these identifications will result in further criminal justice proceedings. Because of the desirability of ecological validity, confidence–accuracy correlations were calculated for only those individuals who identified the suspect. See Table VI for confidence–accuracy correlations for suspect identifications, and see Table X for an examination of the differences between informed and uninformed groups with respect to the effectiveness of accuracy predictors. Accuracy was defined as an identification of the criminal while inaccuracy was defined as an identification of the innocent suspect.

Predecision confidence was not significantly correlated with accuracy for the informed group ($r = .189$, $p = .366$, $n = 25$) or for the uninformed group ($r = -.370$, $p = .131$, $n = 18$). However the correlation for the informed group was marginally higher than the correlation for the uninformed group ($Z = 1.729$, $p = .084$).

Postdecision confidence was significantly correlated with accuracy for the informed group ($r = .529$, $p = .007$, $n = 25$) but not for the uninformed group ($r = .295$, $p = .235$, $n = 18$). However, the postdecision confidence–accuracy correlation for the informed group was not significantly greater than the same correlation for the uninformed group ($Z = .851$, $p = .395$).

Willingness to view a live lineup was not significantly correlated with accuracy for the informed group ($r = .174$, $p = .520$, $n = 16$), but it was correlated with accuracy for the uninformed group ($r = .629$, $p = .029$, $n = 12$). However, the difference between informed and uninformed groups was not significant ($Z = 1.301$, $p = .194$).

Willingness to sign a statement was significantly correlated with accuracy for both the informed group ($r = .719$, $p = .006$, $n = 13$) and the uninformed group ($r = .765$, $p = .010$, $n = 10$). The difference between informed and uninformed groups was not significant ($Z = .207$, $p = .834$).

There are two problems with basing the confidence–accuracy correlations only on suspect identifications. First, only a small percentage of the witnesses identified the suspect, therefore, the correlations are based on a small number of subjects. Second, the criminal justice system sometimes uses multiple suspect lineups thereby increasing the ecological validity of correlations based on all choosers (i.e., witnesses who identify a foil or a suspect). Ecological validity would seem to preclude the inclusion of nonchoosers, however, because witnesses who do not make an identification are not likely to be asked to testify in court (Wells & Lindsay, 1980). Basing the correlations on all witnesses who make an identification (i.e., choosers), allows the experimenter to increase the number of subjects who contribute to the correlations without losing ecological validity. In this case, accuracy was defined as an identification of the criminal while inaccuracy was defined as an identification of the innocent suspect or an identication of a foil. For all witnesses who made a choice, there was a correlation (r = .331, p = .001, n = 93) between the postdecision confidence measure and accuracy. There was also a correlation (r = .405, p = .005, n = 47) between willingnes to sign a statement and accuracy. See Table VII. The accuracy correlations were consistently higher for the informed group. However, the differences between the informed and uninformed groups were not statistically significant. See Table X.

For the uninformed group, both of the behavioroid measures were significantly correlated with accuracy when the correlations were only based on suspect identifications (i.e., criminal or innocent suspect). However, when the correlations were based on all identifications, the behavioroid measures for the uninformed group were not significantly correlated with accuracy. The uninformed group made a large number of foil identifications. The fact that the behavioroid measures were correlated with accuracy for suspect identifications but not for identifications in general indicates that uninformed subjects were more confident of their identifications of the criminal as compared to identifications of the innocent suspect, but not more confident of their identifications of the criminal as compared to their identifications of the foils.

It's possible to argue that non-identifications should also be included in any confidence-accuracy correlation since non-identifications of the innocent suspect are a form of accuracy, while non-identifications of the criminal are a form of inaccuracy. In addition, Wells & Lindsay (1980) have shown that non-identifications provide information, which can be used to revise the predecision probability that the suspect is the criminal. Thus, even though courts must deal almost exclusively with choosers only, decreasing the probability that the suspect is the criminal (i.e., suspect is not identified) can be as important as increasing the probability that the suspect is the criminal (i.e., the suspect is identified). For this reason, confidence-accuracy correlations were also recalculated for nonchoosers only. See Table VIII. For nonchoosers, accuracy was defined as a refusal to make an identification if a criminal absent lineup was being viewed. Inaccuracy was defined as a refusal to make an identification if a criminal present lineup was being viewed.

Since only those subjects who made an identification were asked if they were willing to sign a statement, an analysis of nonchoosers was not possible for this measure. None of the correlations for nonchoosers were significant.

Choosers and nonchoosers were combined in order to provide a total confidence-accuracy picture. See Table IX. For the criminal present conditions, accuracy was defined as an identification of the suspect (i.e., criminal). For the criminal absent conditions, accuracy was defined as a refusal to make an identification. For the criminal present conditions, inaccuracy was defined as an identification of a foil or a refusal to make an identification. For the criminal absent conditions, inaccuracy was defined as an identification of the suspect (i.e., innocent suspect) or an identification of a foil. The predecision confidence-accuracy correlation and the postdecision confidence accuracy correlation were marginally higher for the informed group than for the uninformed group (Zs = 1.694 and 1.942 respectively; p = .091 and .052 respectively. See Table X.

Because accuracy might be better predicted by a combination of predictors than by any particular predictor, the three confidence measures and the one cooperation measure were combined into a multiple regression ananlysis. The multiple regression was significant (r = .526, p = .008, n = 47). This multiple regression was for choosers only because willingness to sign a statement was only asked of choosers. The correlation for the informed group (r = .736, p = .007, n = 22) was marginally higher, (Z = 1.788, p = .073), than the correlation for the uninformed group (r = .3643, p = .56, n = 25). See Tables VI and X.

The Embedded Figures Test was added to the confidence measures as an additional predictor of accuracy and a new multiple regression analysis was calculated (r = .543, p = .045, n = 37). With the addition of the Embedded Figures Test, the multiple regression for the informed groups (r = .837, p =.04, n = 14) became significantly higher (Z = 2.059, p = .039) than the multiple regression for the uninformed group (r = .4048, p = .654, n = 23)

In summary, the confidence-accuracy relationship for the informed group was almost always higher than the same relationship for the uninformed group. See Tables VI through IX. In addition, the multiple regression that included the Embedded Figures Test was a significantly better accuracy predictor for the informed group than for the uninformed group. Also when the correlations were based on the entire subject pool, both the predecision confidence measure and the postdecision confidence measure were marginally better accuracy predictors for the informed group than for the uninformed group. For suspect identifications, predecision confidence was a marginally better accuracy predictor for the informed group than for the uninformed group. The above indicates that accuracy predictors that worked for the informed group were not adequate for predicting the accuracy of the uninformed group.

E. Effect of Information Variable on Willingness to Make an Identification

There was no difference between the informed and uninformed group on willingness to make an identification. The information x choice interaction from the overall Chi square was not significant (Chi Square (2) = 2.494, p = .287). In addition, the number of informed witnesses who refused to make an identification did not differ from the number of uninformed witnesses who refused to make an identification (Z = 1.329, p = .184). The presence-absence x information x choice interaction in the overall Chi Square was significant. However, the choice pattern difference between informed and uninformed witnesses resulted from differences in selections of the suspect and from differences in selections of the foils. The informed group did not differ from the uninformed group in the number of refusals to make an identification under either the criminal present condition (Z = .662, p = .509) or under the criminal absent condition (Z = .847, p = .395).

F. Predecision Confidence vs. Postdecision Confidence as Accuracy Predictors

Correlation coefficients were calculated for both the predecision confidence–accuracy relationship and the postdecision confidence–accuracy relationship. The difference between the correlations was not significant, thereby indicating that the predecision confidence measure is not more closely related to eyewitness accuracy than is the postdecision confidence measure. This result holds regardless of whether the data is analyzed in terms of choosers only or in terms of nonchoosers only or in terms of the entire subject pool (Zs = 1.19, .310, .646; $p > .05$). In addition, none of the confidence or cooperation measures is significantly better than any other of the measures at predicting eyewitness accuracy. See Table XI.

G. Replication of the Functional Size Effect

This experiment partially replicates the functional size effect of Lindsay and Wells (1980).

The size x choice interaction from the overall Chi Square was significant (Chi Square (2) = 6.159, $p = .046$). See Table XIII. Further analysis of this interaction indicated that low functional size resulted in significantly ($Z = 2.470$, $p = .014$) fewer foil selections than did the high functional size lineups. The low and high functional size groups did not differ in selections of the suspect or in refusals to identify someone (Zs = 1.050 & 1.329 respectively; p = .294 & .184 respectively). This indicates that the functional size of the lineup did affect witness choices. Although size affects witness choices, this pattern is not in line with the conceptual expectations regarding the functional size effect as described by Lindsay & Wells (1980) and Wells, Leippe & Ostrom (1979).

The functional size effect is a hypothesis about the change in the ratio of identifications of the criminal relative to identifications of the innocent suspect. This change is hypothesized to occur because similarity among lineup members is increased, thereby, decreasing choices of both the innocent suspect and the criminal. However, the hypothesis predicts that choices of the innocent suspect will decrease faster than choices of the criminal. In order to test this hypotheses, one needs to examine the selections of the criminal relative to selections of the innocent suspect. One way to do this is to look at the presence-absence x size x choice interaction from the overall Chi Square. This interaction tells whether the high functional size condition resulted in a different presence-absence x choice interaction than did the low functional size condition. The three-way interaction was not significant (Chi Square (2) = 1.924, p = .382). Nevertheless, because of the importance of the functional size hypothesis and the fact that there were specific predictions about its effect, specific analyses within the presence-absence x size x choice interaction were conducted.

When the criminal was present the size by choice interaction was marginally significant (Chi Square (2) = 5.712, p = .057). See Table XIV. In addition, when the criminal was present, the suspect (i.e., the criminal) was identified exactly the same number of times by the high functional size group as by the low functional size group (i.e., 16). However, when the criminal was present, the low functional size group selected significantly fewer foils (Z = 2.171, p = .03) than did the high functional size group; and the low functional size group decided marginally more often than did the high functional size group to refuse to make an identification (Z = 1.917, p = .055).

When the criminal was absent, the size x choice interaction was also marginally significant (Chi Square (2) = 4.886, p = .087). See Table XIV. In addition, when the criminal was absent, the number of times the high functional size group refused to make an identification exactly equalled the number of times that the low functional size group refused to make an identification (i.e., 27). As was the case when the criminal was present, the low functional size group choose fewer foils than did the high functional size group. However, when the criminal was absent, the difference in selections of foils was not significant. The criminal absent analysis also revealed that the low functional size group selected the suspect (i.e. innocent suspect) significantly more often than did the high functional size group (Z = 2.02, p = .043).

In summary, when the criminal is present, the high and low functional size groups do not differ in selection of the suspect (i.e., criminal) but the high functional size group did select more foils at the expense of refusals to make an identification. When the criminal was absent, the high and low functional size groups did not differ in refusals to make an identification but the low functional size group identified the innocent suspect more often than did the high functional size group. These results have an important applied implication (i.e., increasing functional size may be able to decrease identifications of the innocent suspect without decreasing identifications of the quilty suspect). However, the results must be viewed with caution because of the nonsignificant presence–absence x size x choice interaction.

Since the functional size hypotheses is a statement about choices of the innocent suspect relative to choices of the criminal there is another possible way of looking for a functional size effect. It's possible to calculate a presence x size Chi Square for only those individuals who identified the suspect (i.e., criminal or innocent suspect). This Chi Square is, of course, based on a much smaller number of subjects. Nevertheless, the Chi Square was marginally significant (Chi Square (1) = 2.8, p = .094). As can be determined from the previous discussion, when the criminal was present, the high functional size group and low functional size group identified the criminal an equal number of times (i.e., 16). When the criminal was absent, the innocent suspect was identified marginally more often (Z = 1.675, p = .093) by the low functional size group. See Table XV.

In summary, there is evidence that functional size of the lineup affects witnesses' choices. However, the evidence for a functional size effect is weak, supported by a presence x size interaction for the suspect identifications that was only marginally significant. Additional support for a functional size effect comes from an internal analysis of the size x presence x choice interaction from the overall Chi Square. This analysis shows that the innocent suspect is selected significantly less often by the high functional size group, while high and low functional size groups identify the criminal at exactly the same rate.

H. Effect of the Information Variable on Functional Size Effect

The information x size x choice interaction from the Chi Square analysis was not significant (Chi Square (2) = 2.498, p = .287). The information x size x choice interaction can be used to examine whether the choices under the low and high functional size conditions differed for the informed and uninformed witnesses (i.e., whether the information variable interacted with the size variable). However, in order to determine whether the functional size effect is affected by the information variable, it is necessary to include the presence-absence variable in the analysis. The presence-absence x information x size x choice interaction from the Chi Square analysis was also not significant (Chi Square (2) = 1.972, p = .373).

As was discussed previously, the functional size hypothesis is a hypothesis about identifications of the suspect. For this reason, a Chi Square was calculated for the subset of individuals who identified the suspect. This Chi Square examined the presence-absence x information x size interaction. This Chi Square was not significant (Chi Square (1) = .248, p = .618). See Table XVI. In summary, there was no evidence that the information variable influenced the functional size effect.

I. Other Findings

Embedded Figures Test

The Embedded Figures Test was not significantly correlated with accuracy. This was true regardless of whether the correlation was based on the entire subject pool or only on those individuals who made an identification of the suspect or an identification of a foil (i.e., choosers) (r's = .091 and .151 respectively; p = .266 and .192 respectively).

Intercorrelations between the three confidence measures and the cooperation measure

The three confidence measures and the cooperation measure tended to be highly correlated with one another. See Table XII for the correlations between confidence measures.

III. Discussion

The information variable affected the distribution of witnesess' identification errors. When the criminal was present in the lineup, uninformed witnesses identified foils more often and identified the suspect less often than did the informed witnesses. No differences between identifications of informed and uninformed witnesses occurred in criminal absent conditions. In addition, there is evidence that the information variable affects the accuracy–confidence relationship in such a manner that the use of confidence as a predictor of accuracy may work less well in actual criminal identifications than it does when the witnesses know that they are participating in an experiment. The information variable does not appear to have an effect on confidence of witnesses or on witnesses' willingness to make an identification or on the functional size effect.

A. Effect of Information Variable on Accuracy

The original suggestion that the information variable might affect accuracy was based on two separate hypotheses. The first hypothesis suggested that accuracy might decline for the uninformed group since at recall they would be operating at a higher level of arousal.

The presence–absence x information x choice interaction in the overall Chi Square can not be explained in terms of the arousal hypothesis. The uninformed witnesses were only less accurate than the informed witnesses when the criminal was present. When the criminal was absent, accuracy was not affected by the information variable. A decrement due to increased arousal at retrieval should have the same effect on both the criminal present and criminal absent conditions. The explanation also can not be in terms of the uninformed group tending to refuse to make identifications and, thereby, automatically being accurate in the criminal absent conditions. Informed and uninformed groups did not significantly differ in refusals to make an identification. In addition, when the criminal present condition and criminal absent condition were analyzed separately, informed and uninformed witnesses still did not significantly differ in refusals to make an identification.

The presence–absence x information x choice interaction indicates that when the criminal was present informed and uninformed witnesses did not differ with respect to whether or not an identification was made. However, when the criminal was present, the uninformed group made fewer identifications of the suspect and more identifications of the foils than did the informed group. When the criminal was absent, informed and uninformed witnesses did not differ in their pattern of choices.

The second hypothesis stated that the consequences of an identification are different for the real world witness than for the laboratory witness. It was suggested that this difference might cause a criterion shift, which in turn, might result in increased accuracy for the uninformed group. The differential consequences hypothesis is equally untenable. That is, the uninformed group was less accurate than was the informed group. In addition, the reduction in accuracy was not the result of the uninformed group being less willing to commit themselves to a choice. When the criminal is present, the uninformed subjects behaved as though they believed he was present and made a selection. Unfortunately, they often selected a foil. When the criminal was absent, the uninformed subjects behaved as though the criminal was absent and increased their refusals to make an identification. This type of behavior can not be explained in terms of increased caution. If the witnesses were merely being cautious, a refusal to identify would have been the most appropriate response for both the criminal present and criminal absent conditions. Instead, the uninformed, criminal present witnesses identified a foil. This is particularly important since our witnesses believed that every lineup member was a suspect. Therefore, an identification of a foil would not be harmless. Any explanation for the witnesses' behavior must include an explanation for increased identifications of the foils when the criminal was present.

When the criminal present lineups were used, the uninformed group was less accurate than when the criminal absent lineups were used. They were also less accurate than their counterparts in the criminal present informed groups. This indicates that the information variable interacts with possibly the most important variable in the eyewitness paradigm (i.e., the presence–absence variable).

Nevertheless, this experiment does not support either of the initial hypotheses.

The data show an unexpected and pronounced tendency for Foil 2 to be 2-1/2 times more likely to be identified when the thief was present than when the thief was absent in the uninformed, high functional size conditions. This runs counter to all previous literature and seems to defy the logic of the situation. It seems only reasonable that the presence of the target should have diminished rather than enhanced the rate at which distractors were chosen. At this point there seems to be no logical explanation for which there is firm data.

Nevertheless, the information variable interacts with at least one variable (i.e., the presence-absence variable) in the eyewitness paradigm. With respect to choices of foils and choices of the suspect, the informed and uninformed witnesses had opposite choice patterns when the criminal was present but not when the criminal was absent. If the real world lineup happened to be a criminal present lineup, there is evidence that real world witnesses who made an idenfication would not only behave differently from informed laboratory witnesses but the real world witnesses would behave in a manner that was totally unpredicted by the behavior of informed witnesses. On the other hand, if the criminal was not in the real world lineup, there is evidence that laboratory witnesses and real world witnesses would behave similarly. However, in the real world it isn't possible to distinguish a criminal present lineup from a criminal absent lineup. Therefore, it isn't possible to generalize from informed eyewitnesses to real world eyewitnesses since it's always possible that the generalization is being made to a criminal present lineup.

In an applied sense, this experiment indicates that one should only have one suspect in a lineup. If the criminal's presence increases the rate at which uninformed witnesses select foils, it is imperative that two suspects, one of whom may be innocent, never appear in the same lineup.

B. Effect of Information Variable on Confidence Ratings and on Predictors of Accuracy

The informed and uninformed groups did not differ in the amount of expressed confidence. However, the information variable affected the relationship between confidence and accuracy. In general, the informed group tended to have higher but not significantly higher, correlations between the predictors of accuracy and accuracy. Overall, informed and uninformed witnesses marginally differed on both the predecision confidence-accuracy correlation and the postdecision confidence-accuracy correlation. When the correlations were based on only those individuals who identified a suspect (i.e., criminal or innocent suspect), informed and uninformed subjects still marginally differed with respect to the predecision confidence-accuracy correlation. In addition, the difference between the two groups was marginally significant when the three confidence measures and one cooperation measure were combined into a multiple regression. More importantly, when the three confidence measures and the cooperation measure and the Embedded Figures Test score were all combined into a multiple regression, the differences between informed and uninformed groups became significant. For all of the marginally significant and significant comparisons, the informed group's confidence-accuracy correlation was superior to the uninformed group's confidence-accuracy correlation.

The results indicate that it's possible to develop a combination of predictors for informed witnesses which, when combined into a multiple regression, account for 70% of the variance. Unfortunately, this same set of predictors can only account for 16% of the variance of the uninformed witnesses. The information variable affected witness accuracy but did not affect witness confidence. This is precisely the type of phenomenon that Leippe (1980) described and that led to his prediction that accuracy-confidence relationships would decrease under real world conditions. Therefore, it is not surprising that the accuracy-confidence relationship is affected by the information variable. By relying on a combination of predictors, which work in the laboratory, one may be mislead into believing that the combination will predict accuracy in the real world. However, the combination may be totally inadequate for separating accurate and inaccurate witnesses in the real world. It is already known that witnesses are not sensitive to situational variables, which affect accuracy (Lindsay et al., 1980). Variables operating in the real world may be another example of situational variables that affect witnesses' accuracy without an appropriate modification of witnesses' confidence.

C. Willingness to Make an Identification

No differences in willingness to make an identification may appear to contrast with Malpass and Devine (1980) in which increasing crime seriousness increased the subjects' willingness to make identifications. However, Malpass and Devine explain their increased willingness in terms of the witness's personal cost-benefit analysis. Increasing crime seriousness served to justify the personal cost (e.g., possible court appearance) for the witness. In the present study, the personal cost to the low seriousness group (informed group) is the same whether the subject makes an identification or not. Considerations faced by Malpass and Devine's low seriousness group do not apply to the present low seriousness group. All of Malpass and Devine's (1980) witnesses were uninformed witnesses (i.e., believed that a crime had occurred). Seriousness was manipulated by increasing the consequences for the criminal. In the present study, it might be argued that seriousness was manipulated by manipulating the information variable (i.e., half the subjects believed that a crime had occurred, while the other half believed that the crime was simulated). If we consider the information variable to be a manipulation of seriousness, it is at least a different type of seriousness than was the case in Malpass and Devine. In this study, it might be argued that the uninformed group may have been motivated to make an identification (i.e., equivalent to Malpass and Devine's high serious group). However, since the informed group was not faced with the negative personal side effects of a real world eyewitness identification, they may have been as willing as the uninformed group to attempt an identification.

D. Predecision Confidence vs. Postdecision Confidence Ratings

Although this study did not find that the predecision confidence measure was more closely related to accuracy than the standard postdecision confidence measure, it may be that other predecision measures would fare better. This study measured confidence prior to seeing the lineup. One could possibly measure confidence after viewing the lineup, but before making a decision. One might expect that it is difficult to estimate confidence when one has not yet seen the pictures, but, easier to estimate confidence after seeing the pictures. However, once the pictures have been seen, the witness may automatically make a covert identification decision. thereby making this measure a type of postdecision measure.

E. Functional Size Effect

Functional size is a hypotheses about suspect identifications; and therefore, the most appropriate method for analyzing a possible functional size effect is the presence-absence x size interaction for suspect identifications. This Chi Square was only marginally significant. However, there were relatively few suspect identifications; and so this Chi Square is based on a small number of subjects (i.e., 42 subjects or 23% of the total number of subjects). It's interesting to note that when the criminal was present, witnesses in the high and low functional size conditions identified the criminal equally often. All errors in the criminal present conditions occurred because the foils drew choices at the expense of the refusal to make an identification catagory. (This can be determined from Table III and from Table XIV). In the real world, if a one suspect lineup is used, identifications of a foil do not result in prosecution of an innocent suspect. The foil is known to be innocent and no prosecution ensues. That is, in the real world criminal present lineup, choices of a foil and refusals to make an identification are functionally equivalent.

On the other hand, when a criminal absent lineup was used, the high functional size lineup reduced the selection of the innocent suspect. In addition, when the criminal was absent, the correct response is a refusal to make an identification. Both Table III and Table XIV indicate that high and low functional size witnesses choose this response equally often.

When the criminal was present, increasing functional size did not detract from identifications of the criminal. When the criminal was absent, increasing functional size did not reduce refusals to make an identification. However, the high functional size did provide a reasonable alternative to the innocent suspect, and thereby, reduced false identifications. The low functional size lineup increased the probability of a false identification of the innocent suspect because it increased the probability that a witness would identify the innocent suspect if the witness merely identified the person that looked most like the criminal. From an applied perspective, these results suggest that high functional size lineups may result in negligible cost (i.e., nonidentifications of the criminal) to the criminal justice system while simultaneously reducing identifications of the innocent suspect.

F. Effect of Information Variable on Functional Size Effect

It was originally speculated that the information variable might interact with functional size. The fact that only a marginal functional size effect was obtained, makes it difficult to know what relationship, if any, exists between the information variable and the functional size variable.

It's possible that a functional size effect difference between informed and uninformed groups was artificially hidden by the experimental procedure. All subjects in this experiment were treated atypically for eyewitness research. All subjects experienced accusations and questioning from the experimenter. They anticipated an interaction in the near future, with Campus Security, and they observed the experimenter's upset over the equipment loss. All of these things were experienced before the informed group discovered that the theft was an experiment. The subjects became emotionally upset over the theft and the upcoming identification.

It may be that this experience made the informed group particularly careful to avoid a second deception by us. Many of the subjects mentioned being chagrined at being deceived by us since they had previously been deceived in one or two other psychology experiments. In addition, subjects' judgments regarding the thief's personality (e.g., attractiveness), may have simplified the identification task. When a criminal present, low functional size lineup was used, the now cautious informed subjects may have thought the task was too easy and suspected a further trick. This suspicion may have caused the subjects to play it safe and refuse to make an identification. When a criminal present, high functional size lineup was used, the task difficulty may have been sufficient to prevent suspicion of a possible second deception.

There is some support in the data that this, in fact did happen. The informed, criminal present, high functional size subjects identified the criminal 50% more often than did the informed, criminal present, low functional size subjects (i.e., 12 identifications vs. 8). While this difference was not statistically significant ($Z = 1.16$, $p = .123$), the result is unusual. This is the only instance in either this experiment or in Lindsay & Wells (1980) where the high functional size group identified the criminal more often than did the low functional size group. It is also surprising because identifying the criminal from the high functional size lineup is believed to be more difficult than identifying the criminal from the low functional size lineup. Increased difficulty is thought to occur because the high functional size lineup contains more feasible distractors, and these distractors draw choices away from the suspect. If the experimental procedure reduced the robustness of the functional size effect for informed subjects, this could not only explain why informed and uninformed subjects did not differ with respect to the functional size effect but it could also explain why the overall functional size effect was only marginally significant.

The inclusion of a third group would have been beneficial. This third group should have received the standard eyewitness experience (i.e., immediate information that the theft was staged). This would have allowed the comparison of the uninformed group to the standard informed group and to an informed group, which only differed from the uninformed group on the information variable.

G. Summary

In addition to the information variable, this experiment manipulated two variables--presence-absence of the criminal in the lineup and functional size. Possible predictors of accuracy were also assessed. The information variable interacted with the presence-absence variable. In addition, it was demonstrated that there exists at least one combination of accuracy predictors, which was a powerful predictor for informed subjects (i.e., accounted for 70% of the variance), but not for uninformed subjects (i.e., accounted for 16% of the variance). The functional size effect was only marginally significant, and the information variable did not signficiantly modify the functional size effect. However, it is possible that the experimental design obscurred both the functional size effect and differences between informed and uninformed witnesses with respect to the functional size effect. This study indicates caution should be used in generalizing from the laboratory to the real world. In addition, the inclusion of a third group, which receives the standard eyewitness experience, would make future research easier to interpret.

Table I

Matrix Depicting Situation for Witnesses in the Real World

	Identification	No Identification
Criminal Present	Box 1 Hit You appear in court. Criminal possibly goes to jail.	Box2 Miss Criminal is released.
Criminal Absent	Box 3 False Alarm You appear in court. Innocent person possibly goes to jail.	Box 4 Correct Rejection Innocent person is released.

Table II

Matrix Depicting Situation for Witnesses in the Eyewitness Laboratory

	Identification	No Identification
Criminal Present	Box 1 Hit Satisfaction of a correct choice.	Box2 Miss Dissatisfaction of an error.
Criminal Absent	Box 3 False Alarm Dissatisfaction of an error.	Box 4 Correct Rejection Satisfaction of a correct choice.

Table III

Breakdown of Choices by Experimental Condition

	Criminal	Innocent Suspect	Foil 1	Foil 2	Foil 3	Foil 4	Foil 5	All Foils	None
Informed									
High Functional Size									
Criminal Present	.52	n/a	0	.13	.04	0	0	.17	.30
Criminal Absent	n/a	.04	.04	.13	.17	0	.04	.39	.57
Low Functional Size									
Criminal Present	.35	n/a	0	.09	0	0	0	.09	.57
Criminal Absent	n/a	.17	.04	.13	.09	.04	0	.30	.52
Uninformed									
High Functional Size									
Criminal Present	.17	n/a	0	.43	0	.04	.04	.52	.30
Criminal Absent	n/a	.04	0	.17	.04	0	.13	.35	.61
Low Functional Size									
Criminal Present	.35	n/a	.04	.04	.09	.04	0	.22	.43
Criminal Absent	n/a	.17	.04	.04	.09	0	0	.17	.65

Table IV

Chi Square Analysis

Source of Variation	DF	Chi Square*	Significance of Chi Square
Total	2	36.712	.000
2-Way Interactions			
Presence x Choice	2	15.190	.001
Information x Choice	2	2.494	.287
Size x Choice	2	6.159	.046
3-Way Interactions			
Presence x Information x Choice	2	6.475	.039
Presence x Size x Choice	2	1.924	.382
Information x Size x Choice	2	2.498	.287
4-Way Interactions			
Presence x Information x Size x Choice	2	1.972	.373

*Note: Chi Square does not permit one to include all subjects and examine how the presence-absence variable interacts with the size variable. Any analysis that includes all subjects and does not include the choice variable is *meaningless. This is true because the experimenter determines how many subjects will be in each catagory except for* the choice catagory. In a case, like the present one, where an equal number of subjects are assigned to each of the experimentally *created catagories, any interaction that involves only the presence-absence, information, and/or size* catagories will necessarily be zero.

Table V

Presence-Absence x Information x Choice Interaction From Chi Square Analysis

	Criminal Present Condition Suspect Identifications	Foil Identifications	Refusals to Identify
Informed Group	20	6	20
Uninformed Group	12	17	17

Note: Informed and uninformed subjects marginally differed with respect to suspect identifications in the criminal present condition.

Note: Informed and uninformed subjects significantly differed with respect to foil identifications in the criminal present condition.

	Criminal Absent Condition Suspect Identifications	Foil Identifications	Refusals to Identify
Informed Group	5	16	25
Uninformed Group	5	12	29

All subjects - Presence-Absence x Information x Choice Chi Square (2) = 6.475, p = .039

Criminal Present Subjects - Information x Choice Chi Square (2) = 7.504, p = .023

Criminal Absent Subjects - Information x Choice Chi Square (2) = .868, p = .648

Table VI

Predictors of Accuracy

Suspect Identifications Only

Predictor*	Informed & Uninformed Groups			Informed Group			Uninformed Group		
	Number of Subjects	r	p	Number of Subjects	r	p	Number of Subjects	r	p
I	43	.082	.603	25	.189	.366	18	-.370	.131
II	43	.401	.008	25	.529	.007	18	.295	.235
III	28	.432	.022	16	.174	.520	12	.629	.029
IV	23	.749	.000	13	.719	.006	10	.765	.010

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement

Table VII

Predictors of Accuracy

Choosers Only

Predictor*	All Choosers Number of Subjects	r	p	Informed Group Number of Subjects	r	p	Uninformed Group Number of Subjects	r	p
I	93	.147	.167	47	.272	.064	46	.017	.909
II	93	.331	.001	47	.459	.001	46	.251	.093
III	53	.148	.290	25	.152	.470	28	.165	.402
IV	47	.405	.005	22	.500	.018	25	.316	.124
V	47	.526	.008	22	.736	.007	25	.364	.560
VI	37	.543	.045	14	.837	.040	22	.405	.654

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement
*V = Multiple regression on I through IV
*VI = Multiple regression on I through IV plus Embedded Figures Test

Table VIII

Predictors of Accuracy

Nonchoosers Only

Predictor*	All Nonchoosers: Number of Subjects	r	p	Informed Group: Number of Subjects	r	p	Uninformed Group: Number of Subjects	r	p
I	91	.004	.969	45	.083	.590	46	-.094	.534
II	90	-.043	.690	45	.100	.512	45	-.173	.255
III	65	.008	.951	35	-.069	.695	30	.150	.430

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup

Table IX

Predictors of Accuracy

All Subjects (i.e., Choosers and Nonchoosers)

Predictor*	All Subjects			Informed Group			Uninformed Group		
	Number of Subjects	r	p	Number of Subjects	r	p	Number of Subjects	r	p
I	184	.049	.505	92	.166	.113	92	-.086	.414
II	183	.143	.054	92	.280	.007	91	-.004	.973
III	118	.042	.652	60	.022	.870	58	.041	.761
IV	47	.405	.005	22	.500	.018	25	.316	.124

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement

Table X

Difference Between Informed and Uninformed Groups on Accuracy Predictors

	All Subjects		All Choosers		Suspect Ids		Nonchoosers	
Predictor*	Z	p	Z	p	Z	p	Z	p
I	1.694	.091	1.222	.222	1.729	.084	.816	.412
II	1.942	.052	1.119	.263	.851	.395	1.260	.207
III	.101	.920	.048	.960	1.301	.194	.842	.401
IV	.709	.478	.709	.478	.207	.834		
V			1.788	.073				
VI			2.059	.039				

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement
*V = Multiple regression on I through IV
*VI = Multiple regression on I through IV plus Embedded Figures Test

Table XI

Comparison of Confidence and Cooperation Measures

To Determine Whether Any One Measure is a Better Predictor of Accuracy

Measure*	Z	p
II–I	1.337	.180
III–I	.006	.992
IV–I	1.533	.126
II–III	1.106	.267
IV–II	.468	.638
IV–III	1.359	.174

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement

Table XII

Correlations Between the Confidence and Cooperation Measures

Measure*	r	n	p
I with II	.3510	183	.000
I with III	.0300	118	.747
I with IV	.1013	47	.498
II with III	.3296	118	.000
II with IV	.4057	47	.003
III with IV	.5832	47	.000

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement
*V = Multiple regression on I through IV
*VI = Multiple regression on I through IV plus Embedded Figures Test

Table XIII

Size x Choice Interaction From Chi Square Analysis

	Suspect Identifications	Foil Identifications	Refusals to Identify
High Size	18	33	41
Low Size	24	18	50

Note: Subjects in the high and low size groups significantly differed with
respect to identifications of the foils.

All subjects – Size x Choice Chi Square (2) = 6.159, p = .046

Table XIV

Presence–Absence x Size x Choice Interaction
From Chi Square Analysis

Criminal Present Condition

	Suspect Identifications	Foil Identifications	Refusals to Identify
High Size	16	16	14
Low Size	16	7	23

Note: High and low functional size subjects significantly differed with respect to foil identifications in the criminal present condition.

Note: High and low functional size subjects marginally differed with respect to refusals to make an identification in the criminal present condition.

Criminal Absent Condition

	Suspect Identifications	Foil Identifications	Refusals to Identify
High Size	2	17	27
Low Size	8	11	27

Note: High and low functional size subjects significantly differed with respect to identifications of the suspect in the criminal absent condition.

All subjects – Presence–Absence x Size x Choice Chi Square (2) = 1.924, p = .382

Criminal Present Subjects – Size x Choice Chi Square (2) = 5.712, p = .057

Criminal Absent Subjects – Size x Choice Chi Square (2) = 4.886, p = .087

Table XV

Suspect Identifications Only

Presence-Absence x Size Interaction

	Criminal Present	Criminal Absent
High Size	16	2
Low Size	16	8

Note: High and low functional size subjects marginally differed with respect to identifications of the suspect when the criminal was absent.

Suspect Identifications Only - Presence-Absence x Size Interaction Chi Square (1) = 2.8, p = .1

Table XVI

Suspect Identifications Only

Presence-Absence x Information x Size Interaction

Criminal Present Condition

	High Size	Low Size
Informed Group	12	8
Uninformed Group	4	8

Criminal Absent Condition

	High Size	Low Size
Informed Group	1	4
Uninformed Group	1	4

Suspect Identifications Only - Presence-Absence x Information x Size Interaction Chi Square (1) = .248, p = .618

Suspect Identifications Only - Criminal Present Condition - Information x Size
Chi Square (1) = 4, p = .046

Suspect Identifications Only - Criminal Absent Condition - Information x Size
Chi Square (1) = 0, p = 1

Abernathy, E. The effect of changed environmental conditions upon the results of college examinations. *Journal of Psychology,* 1940, *10,* 293-301.

Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology,* Vol. 6, New York: Academic Press, 1972.

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., and Sturgill, W. Memory for faces and the circumstances of encounter. *Journal of Applied Psychology,* 1977, *62,* 311-318.

Clifford, B. R. & Bull, R. *The psychology of person identification.* Boston: Routledge & Kegan Paul, 1978.

Cohen, M. E., & Carr, W. J. Facial recognition and the von Restorff effect. *Bulletin of the Psychonomic Society,* 1975, *6,* 383-384.

Craik, F. I. M. and 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 and Psychophysics,* 1971, *10,* 393-396.

Dion, K. Physical attractiveness and evaluations of children's trangressions. *Journal of Personality and Social Psychology,* 1972, *24,* 207-213.

Dion, K.; Berscheid, E.; and Walster E. What is beautiful is good. *Journal of Personality and Social Psychology,* 1972, *24,* 285-290.

Easterbrook, J. A. The effect of emotion on utilization and organization of behavior. *Psychological Review,* 1959, *66,* 183-201. Egan, J. P. *Signal Detection Theory and ROC Analysis.* New York: Academic Press, 1975.

Fanselow, M. S. & Buckhout, R. Nonverbal cueing as a source of biasing information in eyewitness identification testing. Unpublished manuscript, Brooklyn College, 1976.

Galper, R. E. Recognition of faces in photographic negative. *Psychonomic Science,* 1970, *19*, 207-208.

Gibson, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century Crofts, 1967.

Goldstein, A. G., Johnson, K. S., and Chance, J. Does fluency of face description imply superior face recognition? *Bulletin of the Psychonomic Society*, 1979, *13*, 15-18.

Goldstein, A. G., Stephenson, B., & Chance, J. Face recognition memory: Distribution of false alarms. *Bulletin of the Psychonomic Society*, 1977, *9*, 416-418.

Goodenough, D. R. The role of individual differences in field dependence as a factor in learning and memory. *Psychological Bulletin*, 1976, *83*, 675-694.

Laughery, K. R., Fessler, P. K., Lenorovitz, D. R., & Yoblick, D. A. Time delay and similarity effects in facial recognition. *Journal of Applied Psychology*, 1974, *59*, 490-496.

Laughery, K. R., & Fowler, R. H. Factors affecting facial recognition. University of Houston Mug File Project, Report No. UHMUG-3, 1977.

Leippe, M. R. Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior*, 1980, *4*, in press.

Leippe, M. R., Wells, G. L., and Ostrom, T. M. Crime Seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology*, 1978, [illegible] 345-351.

Light, L. L.; Kayra-Stuart, F. and Hollander, S. Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, 197[illegible], [illegible] 212-228.

Lindsay, R. C. L., Wells, G. L., and Rumpel, C. M. Can people detect eyewitnessidentification accuracy within and across situations? *Journal of Applied Psychology*, 1980, in press.

Lindsay, R. C. L. and Wells, G. L. What price justice? Exploring the relationship of lineup fairness to identification fairness. *Law and Human Behavior*, 1980, *4*, in press.

Lipton, J. P. On the Psychology of eyewitness testimony. *Journal of Applied*

Psychology, 1977, *62*, 90-93.

Loftus, E. F. *Eyewitness Testimony*. Cambridge: Harvard University Press, 1979.

Loftus, E. F. Unconscious transference in eyewitness identifications. *Law & Psychology Review*, 1976, *2*, 93-98.

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

Loftus, E. F. and Zanni, G. Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, *5*, 86-88, 1975.

Malpass, R. S. & Devine, P. G. Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, 1980, in press.

Malpass, R. S. and Devine, P. G. Realism and eyewitness identification research. *Law & Human Behavior*, 1980, *4*, in press.

Münsterberg, H. *On the witness stand: Essays on psychology and crime*. New York: Clark, Boardman, 1908.

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

Ross, L. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (ed.), *Advances in experimental social psychology*, *10*, New York: Academic Press, 1977.

Shepherd, J. W., & Ellis, H. D. The effect of attractiveness on recognition memory for faces. *American Journal of Psychology*, 1973, *86*, 627-633.

Sigall, H. and Ostrove, N. Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on judicial judgment. *Journal of Personality and Social Psychology*, 1975, *31*, 410-414.

Sorce, J. F., & Campos, J. J. The role of expression in the recognition of a face. *American Journal of Psychology*, 1974, *87*, 71-82.

Strnad, B. N., & Mueller, J. H. Levels of processing in facial recognition memory. *Bulletin of the Psychonomic Society*, 1974, *9*, 17-18.

Warrington, E. K., & Ackroyd, C. The effect of orienting tasks on recognition memory. *Memory and Cognition*, 1975, *3*, 140-142.

Wells, G. L. Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 1978, *36*, 1546-1557.

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

Wells, G. L. and Lindsay, R. C. L. On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, 1980, *88*, 776-784.

Wells, G. L., Lindsay, R. C. L., and Ferguson, T. J. Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology*, 1979, *63*, 440-448.

Wells, G. L., Lindsay, R. L. C., and Tousignant, J. P. Effects of expert psychological advice on jurors' performance in judging the validity of eyewitness identifications. *Law and Human Behavior*, 1980, in press.

Winograd, E. Recognition memory for faces following nine different judgments. *Bulletin of the Psychonomic Society*, 1976, *8*, 419-421.

Witken, H. A., Oltman, P. K., Raskin, E., and Karp, S. A. *A Manuel for the Embedded Figures Tests.* Palo Alto: Consulting Psychologists Press, 1971.

Yarmey, A. D. The effects of attractiveness, feature saliency, and liking on memory for faces. In M. Cook & G. Wilson (Eds.), *Love and attraction: An international conference*. Toronto: Pergamon Press, pp. 51-53, 1978.

Yarmey, A. D. *The psychology of eyewitness testimony*. New York: Free Press, 1979.

Appendix A

Description of Thief Provided to Mock Witnesses

General Description:	Handsome
Hair:	Brown, parted, straight, about 3-1/2 inches long
Age:	Between 18-22
Weight:	Unknown
Eyes:	Grey or blue, maybe brown
Jaw:	Angular
Complexion:	Clear
Facial Hair:	None

Appendix B

*Criteria for Rating Confidence

Measure*	1	2	3	4	5
I	Not really Definitely not No	Probably not Not likely Not to confident Not sure I don't think so I don't know	I'll try I could try Possibly Maybe	Yeah, I think so Probably, yeah Probably Not positive I think so	If I saw him again Yeah You bet ya Yes
II	I didn't get a good look I didn't look at him Don't know Not sure at all	Not even 50% Not very sure Not sure	Maybe Possibly 50-50	Not 100% sure Not positive Pretty sure 75% sure Not exactly positive If anyone, it's him Can't be positive Fairly sure	Quite sure Sure 98% sure Almost positive Very sure
III	Wouldn't because an id might convict *NO* Definitely not	Wouldn't help Didn't see him No I doubt it	Maybe	Probably Yes, but it might not help I guess	Yes Sure Definitely Certainly
IV	*NO* Definitely not	Doubt it No	Not sure Don't know	Maybe If I could be sure	Yes Definitely Sure

*I = "Do you think you can make an identification?" (i.e., Predecision confidence).
*II = "How sure are you that this is the man?" (i.e., Postdecision confidence).
*III = Willingness to view a live lineup
*IV = Willingness to sign a statement