Distinguishing Accurate From Inaccurate Eyewitness Identifications via Inquiries About Decision Processes

David Dunning and Lisa Beth Stern

Accurate eyewitness identifications can be distinguished from erroneous ones in part by asking witnesses to describe the decision processes that led to their judgments. In 4 studies, mock witnesses were presented with a videotaped staged crime and then asked to identify the perpetrator from a photo lineup. Of those making positive identifications, accurate witnesses were more likely than their inaccurate counterparts to state that their judgments resulted from automatic recognition (e.g., "His face just ‘popped out’ at me"). Inaccurate witnesses more frequently stated that they followed a process of elimination strategy (e.g., "I compared the photos to each other to narrow the choices"). A 5th study revealed that people have partial understanding of these principles, and thus are modestly successful in differentiating accurate from inaccurate identifications. Explicitly informing them of these principles enhanced their discrimination ability even further.

In a drama that is repeated daily in jurisdictions across the United States and elsewhere, witnesses to crimes are asked to come to police stations, to view lineups, and to identify the perpetrator from the choice of individuals given. In many types of cases, such as sexual assault or robbery, eyewitness identifications may be the only evidence ever gathered in crime investigations. Indeed, Goldstein, Chance, and Schneller (1989) discovered that 78,000 trials in the United States during 1987 were decided predominantly on the basis of eyewitness identification evidence. As such, a defendant's fate often rests on whether criminal justice officials or jurors believe the witness has made an accurate identification.

Given this state of affairs, psychological literature on eyewitness testimony paints a disturbing portrait. Eyewitnesses identifications are often convincing (Devlin, 1976), yet many factors can conspire to make eyewitnesses reach erroneous identifications (for reviews, see Loftus, 1979; Ross, Read, & Toglia, 1994; Wells & Loftus, 1984). More troubling, mock jurors have little ability to discern accurate eyewitnesses from erroneous ones (Brigham & Bothwell, 1983; Cutler, Penrod, & Stuve, 1988; Kassin, Rigby, & Castillo, 1991; Leippe, Manion, & Romanczyk, 1992; Lindsay, Wells, & O'Connor, 1989; Wells, Lindsay, & Ferguson, 1979).

These circumstances have led many psychological researchers to search for factors related to accuracy and error in eyewitness identification. To a considerable extent, these researchers have been successful, identifying a number of situational variables that enhance or detract from eyewitness accuracy (for reviews, see Kassin, Ellsworth, & Smith, 1989; Loftus, 1979; Ross, Read, & Toglia, 1994; Wells, 1993; Wells & Loftus, 1984). Witnesses are more likely to be accurate the longer they view the perpetrator's appearance (Laughery, Alexander, & Lane, 1971), and the shorter the delay between crime and identification attempt (Lipton, 1977). Witnesses are also more likely to be accurate when reminded of the context of the crime (Krafka & Penrod, 1985). They are less likely to be accurate if a weapon or dangerous object was present at the crime scene (Steblay, 1992), or if they are asked to identify an individual of a different race from their own (Brigham & Malpass, 1985).

Psycholegal research has lagged in creating more direct tests of eyewitness accuracy, questions that can be directly asked of the witness at the time of the identification to determine whether he or she has correctly identified the culprit or has mistakenly incriminated an innocent individual (Dunning, 1989). To be sure, researchers have looked, but have been largely unsuccessful. For example, an eyewitness's confidence in his or her identification should be related to his or her accuracy, yet many studies have repeatedly shown that confidence is generally unrelated to the veracity of an eyewitness's identification (Bothwell, Deffenbacher, & Brigham, 1987; Smith, Kassin, & Ellsworth, 1989; Wells & Murray, 1984).

In the research contained in this article, we describe one attempt to find a test of eyewitness accuracy that could be applied to specific witnesses at the time they make their identifications. Guided by work on reality monitoring, facial recognition, and eyewitness identification, we proposed that witnesses making accurate versus inaccurate positive identifications would differ in the cognitive processes they report when rendering their judgments. We predicted that accurate witnesses would reach their decisions with relatively little conscious strategy. Thus, when
asked to describe the process leading to their identifications, they would describe their experiences as "effortless" or, more specifically, involving automatic recognition (e.g., "His face just popped out at me"). Inaccurate witnesses, on the other hand, would report that they pursued explicit, deliberative procedures when reaching their judgments, usually expressed in the form of process of elimination (e.g., "I compared the photos to each other to narrow the choices"). Consequently, we would be able to differentiate accurate witnesses from erroneous ones by asking witnesses to describe their decision processes just after they made their identifications.

Suggestions From Work on Reality Monitoring

The assertion that accurate and inaccurate eyewitnesses would describe different cognitive operations is supported by work on reality monitoring (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981). In this work, researchers assess how people distinguish real memories, that is, recollections of perceptually experienced events, from imagined ones. According to the theory, people accomplish this task by noting specific characteristics of their memories. Memories about external events contain a great deal of contextual information about time and place, as well as sensory and perceptual detail (color and shape of objects). In contrast, memories for internal events, produced by means of imagination and fantasy, contain information about the intentional cognitive operations pursued in generating them.

Many studies have shown that external (perceived) and internal (imagined) memories differ on these proposed dimensions. Subjects asked to describe real versus imagined autobiographical events in the recent past rate real events as containing more perceptual information, such as sound, smell, touch, location, setting, and visual detail (Johnson, Foley, Suengas, & Raye, 1988). Subjects are also better able to identify imagined events as internally generated to the extent that producing those events involves deliberative cognitive work (Johnson, Raye, Foley, & Raye, 1981). When information about cognitive effort is missing, as it is for dreams, people have a difficult time distinguishing fantasy from fact (Johnson, Kahan, & Raye, 1984).

These cues can be used to distinguish between real and imagined memories of objects present at the scene of staged crimes. For example, Schooler, Gerhard, and Loftus (1986) asked subjects to view a staged traffic accident. Some subjects saw the offending automobile pass by a yield sign. Others saw a version of the accident in which no yield sign was present, but had the presence of one suggested to them by misleading questions. Later, all subjects were asked to describe the sign. Consistent with reality monitoring principles, subjects who actually saw the yield sign provided more sensory and perceptual information in their recollections than did subjects who were misled and thus only imagined the yield sign. In contrast, misled subjects tended to incorporate a greater amount of cognitive work in their accounts. That is, they tended to describe what they were thinking or paying attention to when viewing the crime, and to mention the functions the yield sign served.

Application to Facial Recognition

How can these differences between real and imagined memories be used to differentiate accurate from inaccurate facial recognition? The situation presented by Schooler et al. (1986) differs substantially from the one confronting witnesses making eyewitness identifications. Schooler et al. found differences in the manner in which subjects described their recollection of an object. However, eyewitnesses are required to recognize the face before them, without offering any description at all. How, then, can accurate and inaccurate acts of recognition be distinguished?

Treatments of reality monitoring suggest an answer. The act of recognition is deciding that an object has been previously encountered in one's external environment. According to reality monitoring principles, this occurs when viewing an object activates memory traces containing perceptual detail (Johnson et al., 1993). To the extent that such perceptual detail is evoked rapidly, the individual will classify the object as externally encountered. When such sensory information fails to be activated, the individual should not recognize the object as previously encountered, although, importantly, mistakes do happen.

The analysis presents a paradox, however, when recognition of faces is involved. When people accurately recognize a face, because the perceptual detail present in the face is a good match to the sensory information contained in their memory, they will be unable to articulate or verbalize what it is about the face or their memory that led them to their reaction. They will act like the subjects in the study by Hay, Young, and Ellis (1986), who were presented with faces of celebrities and then asked to decide whether these faces were present in arrays containing several stimulus faces. Often, these subjects remarked that the relevant faces "just popped out" at them.

This paradox concerning accuracy and articulation arises for two reasons. First, people are experts at facial recognition (Bahrick, Bahrick, & Wittlinger, 1975; Ellis, 1981). They use rapid, automated cognitive operations when recognizing facial stimuli, completed without conscious awareness or monitoring, that leave them unable to articulate the bases of their decisions (Ellis, 1984). Second, people tend to construct representations of faces that contain a good deal of configural information, that is, information concerning the relations and proportions of facial features to each other as well as data about general face shape (Diamond & Carey, 1986; Rhodes, Tan, Brake, & Taylor, 1989; Sergent, 1984; Wells & Hryciw, 1984; Wells & Turtle, 1987; Young, Hellawell, & Hay, 1987). Information about such configural relations is not easily articulated (Fallshoer & Schooler, 1993; Schooler & Engstler-Schooler, 1990).

Several strands of research also suggest that accurate facial recognition tends to be a rapid, automated process. First, in one study of eyewitness identification, Sporer (1993) discovered that accurate positive identifications tended to be made more quickly than inaccurate ones. He also found a weak but consistent correlation between reaction time and eyewitness accuracy in a review of the eyewitness literature (Sporer, 1994). Second, in their studies of facial recognition, Hay et al. (1986) discovered that people recognized the faces of celebrities without cognitive deliberation. Indeed, the speed with which subjects rec-
ognized a celebrity's face was unaffected by the number of distracter faces presented along with the target, indicating that the act of recognition had been an automatic one, conducted without analysis of all the choices presented.

Finally, the notion that accurate facial recognition tends to be a rapid, automatic process is supported by Schooler and Engstler-Schloer's (1990) recent work on verbal overshadowing of face memory (see also Fallshore & Schooler, 1993). In their research, they found that they could degrade people's ability to recognize faces by asking them to provide verbal descriptions of those faces. They argued that this impairment occurred because asking people for verbal descriptions caused them to rely on verbal codes of their memories for the faces they had seen as opposed to more veridical visual codes. However, Schooler and Engstler-Schloer discovered that they could impel subjects to rely on these more accurate visual codes by forcing them to make recognition judgments quickly, that is, within 5 s. When speeded in their judgments, subjects providing verbal descriptions were just as accurate in their recognition judgments as those not providing such descriptions.

But what about unsuccessful facial recognition? When people mistakenly report that they have seen a face before, what form will their decision processes take? Work on reality monitoring suggests that people, if they fail to recognize the perpetrator automatically, will use more deliberate, explicit cognitive strategies to reach their decisions (Johnson et al., 1993). They may concentrate on details of the pictures before them or of the memories they possess about the perpetrator. They may begin to devise explicit decision rules to assist them in making their choices. Whatever specific form these deliberative cognitive operations take, they are likely to involve process of elimination. Witnesses will compare lineup alternatives to each other, eliminate the most implausible options, and choose among the rest.

Past work on eyewitness identification has suggested a link between process of elimination strategies and eyewitness error. Specifically, Wells (1993) proposed that inaccurate positive identifications are often the result of relative judgment strategies in which witnesses compare lineup alternatives with each other to determine the best match to their memory of the culprit. In support of this assertion, Wells (1984) exposed a number of subjects to a staged crime and identified a group of subjects who provided verbal descriptions of those faces. They argued that this impairment occurred because asking people for verbal descriptions caused them to rely on verbal codes of their memories for the faces they had seen as opposed to more veridical visual codes. However, Schooler and Engstler-Schloer discovered that they could impel subjects to rely on these more accurate visual codes by forcing them to make recognition judgments quickly, that is, within 5 s. When speeded in their judgments, subjects providing verbal descriptions were just as accurate in their recognition judgments as those not providing such descriptions.

But what about unsuccessful facial recognition? When people mistakenly report that they have seen a face before, what form will their decision processes take? Work on reality monitoring suggests that people, if they fail to recognize the perpetrator automatically, will use more deliberate, explicit cognitive strategies to reach their decisions (Johnson et al., 1993). They may concentrate on details of the pictures before them or of the memories they possess about the perpetrator. They may begin to devise explicit decision rules to assist them in making their choices. Whatever specific form these deliberative cognitive operations take, they are likely to involve process of elimination. Witnesses will compare lineup alternatives to each other, eliminate the most implausible options, and choose among the rest.

Past work on eyewitness identification has suggested a link between process of elimination strategies and eyewitness error. Specifically, Wells (1993) proposed that inaccurate positive identifications are often the result of relative judgment strategies in which witnesses compare lineup alternatives with each other to determine the best match to their memory of the culprit. In support of this assertion, Wells (1984) exposed a number of subjects to a staged crime and identified a group of subject-witnesses who were likely to pursue a relative judgment strategy, these being subjects who selected a photograph in a lineup that did not include the culprit. When shown a second lineup that contained the perpetrator, these subjects were less likely than others to identify the culprit accurately. When shown a second lineup that again failed to contain the perpetrator, they were less likely than other participants to recognize the culprit's absence.

Other work supports a link between process of elimination and eyewitness error. When process of elimination strategies are precluded, the prevalence of eyewitness error decreases. For example, subjects find it difficult to pursue process of elimination or relative judgment strategies when faced with sequential lineups (Lindsay & Wells, 1985), in which witnesses view the faces in the lineup one at a time. For each face, they must decide whether it is the culprit, with no chance to go back and review previous lineup foils. Under this procedure, the rate of false positive identifications is reduced, with no reduction in the rate of accurate judgments (Cutler & Penrod, 1988; Lindsay & Wells, 1985; Sporer, 1993).

Overview of Present Research

In sum, the general principles of reality monitoring research, as well as supportive evidence from facial recognition and eyewitness identification work, suggest that accurate and inaccurate eyewitnesses will distinguish themselves by how they describe their identification decisions. Accurate witnesses should report little conscious strategy, stating instead that they recognized the perpetrator automatically. Inaccurate witnesses, in contrast, should report following deliberate process of elimination strategies.

To test these predictions, we conducted four studies in which mock witnesses viewed a staged crime and then were asked to identify the culprit from a photo spread. Study 1 was designed as a direct test of these notions concerning the decision processes of accurate and inaccurate eyewitnesses. It was predicted that accurate eyewitnesses would be more likely than their erroneous counterparts to report that their judgment stemmed from automatic recognition processes. In contrast, inaccurate witnesses were predicted to be more likely than their accurate peers to articulate process of elimination strategies.

Study 2 replicated this procedure, except that witnesses received biased instructions when viewing the lineup, that is, they were forced, as much as ethically possible, to make positive eyewitness identifications. Study 3 examined whether the relationship between decision processes and eyewitness accuracy was influenced by the decision criteria subjects were asked to adopt. Some witnesses were told that they had to be fairly certain in their decisions before positively identifying an individual in the lineup (stringent criterion condition), and others were told to positively identify a lineup alternative regardless of their confidence (lenient criterion condition).

Finally, in Study 4, we explored subjects' reactions to perpetrator-absent lineups. With such a lineup, any positive identification an individual makes is an error by definition. We predicted that subjects making false positive identifications from perpetrator-absent lineups would describe their decisions as a product of process of elimination strategies, with few reports of automatic recognition. The inclusion of a perpetrator-absent lineup in Study 4 also allowed us to explore the situation of most interest to criminal justice officials. When these officials ask witnesses to view lineups, they are usually interested in identification of only one person: the suspect. The question facing them is whether identifications of the suspect represent correct identifications from a perpetrator-present lineup or erroneous inclusions from a perpetrator-absent one. All other positive identifications do not matter, because they are inclusions of distractor foils known to be innocent. In Study 4, we could explore whether we could distinguish these two situations by comparing the decision processes of witnesses correctly identifying the offender in the perpetrator-present lineup with those of witnesses identifying his replacement in the perpetrator-absent one.

In all four studies, we also tested supplemental predictions.
drawn from the distinction between automatic recognition versus process of elimination. In three of the studies, we assessed how quickly subject-witnesses reached their decisions. This allowed us to potentially replicate the findings of Sporer (1993, 1994), but also to see whether reaction time was related to, and explained by, the cognitive processes subjects reported while making their decisions.

Additional predictions focused on other reports that subjects could make about the identification experience. Because inaccurate witnesses were expected to use process of elimination strategies, we predicted that they would cite the photos as having an impact on their decisions. This contention was assessed in two different ways. First, subjects were asked about the influence of nonchosen photographs in their decisions. It was predicted that accurate witnesses would state that nonchosen photos had little impact on their decisions, whereas inaccurate witnesses would say that the photos did have an impact by making the decision more difficult (e.g., “They confused me”). Second, we asked all subject-witnesses whether their memories or the photographs had a greater influence on their decisions. We expected that accurate witnesses, because their decisions were based on automatic recognition, would be more likely to cite their memories as having more of an impact on their decisions than would their erroneous peers. Inaccurate witnesses, relying on process of elimination strategies, would be more likely to cite the photographs as influential. These questions were also asked because they reflected common themes observed in eyewitness protocols in preliminary studies leading to this work (Stern & Dunning, 1994; see also Stern, 1991). That work also provided some preliminary evidence that accurate and inaccurate witnesses could be distinguished from their responses to such queries.

In addition to these studies, a fifth investigation assessed whether information about witnesses’ decision processes could help interested individuals to discern accurate identifications from inaccurate ones. Toward that end, subjects were given some of the eyewitness protocols collected in the first four studies and asked to decide which ones were accurate and which were not. Some of these subjects were informed of our findings from the first four studies; others completed the task without this information.

Studies 1–4: Method

Because the first four studies are so similar in procedure, they are described together in the Method and Results sections that follow. We first describe Study 1 in detail. For the remaining three studies, we note significant differences in procedures where they occur. In all four studies, mock witnesses viewed a videotaped staged crime and were later asked to identify the perpetrator from a five-person photo spread. Witnesses making positive identifications (or any definitive decision in Study 4), whether correct or incorrect, were asked about the decision processes they had followed in reaching their decisions.

Study 1

Procedure

Subjects were run individually. The experimenter, blind to the hypotheses under study, explained that the study focused on recruiting people to become elementary school teachers. The experimenter stated that he or she would show the subject “some unedited video” made by the Cornell Psychology Department that was to be turned into a recruitment piece aimed at stimulating interest in the teaching profession. Subjects were asked to pay close attention, for they were to be asked for their comments concerning the quality of the video and its potential effects.

Subjects then viewed the video, which contained many scenes filmed at a preschool. The video focused on one female teacher who was seen leading children through many different activities. The subjects saw the teacher lead preschool students through a game of baseball, read the children a story, and supervise them playing in a sandbox, among other activities. Near the end of the video, the teacher mentioned to another staff member that she was going on break. The camera followed her down the hall to a lounge, where she left her purse on a table while she went to some vending machines nearby. A man sat at a table nearby, wearing a white, blue-striped dress shirt and jeans. He was of European heritage, of college age, had an athletic build, and had brown eyes and short brown hair. He looked over at the purse, looked at the teacher, and then looked around the room to see if anyone was watching. He reached into the purse, pulled out the teacher’s wallet, opened it, and removed some money. He then stood up and walked out of the lounge while putting the stolen money in his pocket. In all, the video lasted approximately 3 min, with the crucial crime episode taking 34 s to transpire. The video was shown on a 19-in. (48.26-cm) video monitor.

At this point, the experimenter informed the subject that the experiment in reality concentrated on eyewitness testimony and that the experimenter had a few questions for the subject to answer. The subject then spent approximately 5 to 10 min answering a questionnaire about the scenes leading up to but not including the final crime scene. Subjects were asked, for example, to describe the games that the children had played and the story that the teacher had read to the students.

After subjects had completed this questionnaire, the experimenter told subjects that he or she would be presenting them with a photograph lineup and would ask them to make an identification judgment. In Study 1, that meant identifying one of the photographs, deciding that the perpetrator was not in the lineup, or refusing to make any decision whatsoever. Subjects were instructed that the perpetrator may or may not be in the lineup. Subjects were also instructed to “say out loud what you are thinking or doing, what sorts of things are going on in your head,” and that they would be tape-recorded while they reached their judgments. After subjects indicated that they understood, the experimenter turned the tape recorder on and presented them with the lineup. The lineup contained five photographs, presented simultaneously, one of which was that of the perpetrator. Subjects then made a judgment. To ensure that the experimenter did not bias the decision of the subjects toward any one photograph, experimenters stood behind subjects during this procedure.

It should be noted that subjects were asked to provide “think aloud” protocols, which were tape-recorded for three reasons. First, it was felt that such a procedure would make it easier for subjects to answer questions later about their decision processes. Second, it was believed that tape-recording subject’s thoughts while they were reaching identification judgments would serve as a potentially rich source of data. Finally, tape-recording subjects allowed us to time how long it took them to reach their decisions. Regrettably, in Study 1 the tape recorder malfunctioned repeatedly, preventing us from garnering information about subjects’ thoughts while they were reaching their judgments. Tape recordings during Studies 2–4 were largely successful.

Besides the photograph of the perpetrator, the four lineup foils came from a collection of 47 photographs of college-age men who possessed the same height, build, coloring, and hairstyle of the perpetrator. The similarity of these photos to that of the offender was rated by 76 stu-
dents. The four photographs selected possessed a range of similarity to the offender, from roughly the median of similarity ratings given to the 47 photographs to one that was judged to be quite similar. In the lineup, all five individuals, including the perpetrator, wore identical red sweatshirts and gazed into the camera without emotion (see Ross, Ceci, Dunning, & Toglia, 1994, for more detail).

Dependent Measures

After reaching a final identification judgment, subjects completed a questionnaire containing the crucial dependent measures under study. The questionnaire began by asking subjects to reiterate their identification choice and to express their confidence in their decision on a scale that ranged from 0% to 100%.

Subjects next completed the measure most central to the aims of the study: "How much influence did the other (nonchosen) pictures have on your decision?" Subjects were presented with eight responses and were told they could endorse as many as apply. The first two responses focused on automatic recognition: "I just recognized him, I cannot explain why," and "His face just 'popped out' at me." The next three responses probed for a process of elimination strategy: "I compared the photos to each other in order to narrow the choices," "I first eliminated the ones definitely not him, then chose among the rest," and "He was the closest person to what I remember, but not exactly." The remaining responses focused on other aspects of the identification process: "I matched the image in my head to the picture in front of me," and "I based the judgment on specific facial features (e.g., nose, hair, eyes)." Subjects could also endorse an option described as "other," with an area for explanation. It should be noted that these response options and their specific wordings were chosen because they reflected common themes observed in open-ended responses of subjects who had participated in preliminary studies leading to the present research (Stern, 1991; Stern & Dunning, 1994).

Subjects next answered the following query: "How much influence did the other (nonchosen) pictures have on your decision?" Subjects were presented with four response alternatives. Two indicated that the nonchosen photographs had had a minimal impact: "They helped me to confirm, reinforce my decision after I made it," and "They had little influence on my decision." Two indicated that the other photographs had made the task more difficult: "They confused me; they made the task more difficult," and "They were all so similar that they made me less confident." Subjects could endorse as many of the options as they felt appropriate. Again, these response alternatives were selected because they reflected common themes we had observed in subjects' open-ended comments in preliminary work (Stern & Dunning, 1994).

Subjects next were asked, "What would you say had a greater influence on your decision, the pictures in the lineup, or your memory of the culprit?" Subjects could answer that their memory had a greater impact, that the pictures had, or that they both had about equal impact in the decision. After answering this question, subjects finished the questionnaire by responding to an open-ended item that requested them to describe, in their own words, the strategy they used in making the identification.

After completing the post-identification questionnaire, subjects were told which photo in the lineup was that of the perpetrator. The experimenter then handed them an envelope containing a debriefing form. The form explained the rationale and predictions of the study, and thanked subjects for their participation. To ensure experimenter blindness, subjects were not allowed to discuss the study further with the experimenter after reading this form.

Subjects

Subjects were 129 Cornell University students drawn from intermediate-level psychology and human development courses. All subjects received extra credit toward their course grades for participation. Data from one additional subject were discarded because he claimed to recognize one of the lineup foils as an acquaintance.

Of the subjects who were asked to make an identification in Study 1, 90 (70%) identified a specific photograph as that of the perpetrator. Of these, 62 (69%) chose the correct photograph and were thus termed accurate witnesses. The remaining 28 (31%) chose an incorrect photo, and were hence labeled inaccurate eyewitnesses. The other 39 subjects in this study declined to identify any of the lineup photographs.

Study 2

Procedures and Dependent Measures

Procedures and dependent measures used in Study 2 were identical to those of Study 1, with the following exception. Subjects were never told that the perpetrator might not be in the lineup. In addition, if subjects asked experimenters whether he may be absent, the experimenters reiterated the fact that the subject's task was to pick someone specifically out of the lineup. As in Study 1, the perpetrator photo was always included in the photo spread.

Subjects

Subjects were 45 Cornell University students drawn from intermediate-level communications and human development courses. All subjects received extra credit toward their course grades for participation. Of the 45 subjects in the study, 41 (91%) identified a specific photograph as that of the perpetrator. Of these, 26 (63%) were accurate; 15 (37%) were inaccurate. Of the remaining witnesses, 2 refused to reach a decision and 2 insisted that the perpetrator's photo was not in the lineup.

Study 3

Procedures and Dependent Measures

Procedures and dependent measures used in Study 3 were largely identical to those of Study 1, with the following exception. Subjects could either positively identify one of the photographs as that of the perpetrator or they could reach no decision. In addition, subjects were assigned to two different criteria conditions that diverged in the level of confidence subjects were asked to reach before providing a positive identification. In the high criterion condition, subjects were told that they must be "clearly certain" or "at least 70% confident" in their judgment when making a positive identification. Otherwise, they were to make no decision. Subjects in the low criterion group were asked to make a positive identification "regardless of how confident" they were. They were not allowed to reach no decision.

Subjects

Subjects were 63 Cornell University students drawn from intermediate-level psychology courses. All subjects received extra credit toward their course grades for participation. Of the 63 subjects in the study, 47 (75%) made a positive identification. Of these, 31 (66%) were accurate; 16 (31%) were inaccurate. All remaining witnesses reached no decision.

Study 4

Procedures and Dependent Measures

Procedures and dependent measures were largely identical with those used in Study 1, with the following exception. Whereas most subjects (n = 75) were presented with a perpetrator-present lineup, others (n = 56)
were shown a perpetrator-absent photo spread, in which the perpeler-

tor's photo was replaced by that of a moderately similar-looking foil. We
assigned subjects to perpetrator-present and perpetrator-absent lineup
conditions at random, with the proviso that the perpetrator-present
lineup condition be overrepresented by a ratio of 7:5. We assigned more
people to the perpetrator-present lineup condition because this circum-
stance produced both accurate and inaccurate witnesses, and we wished
to compare the responses of inaccurate witnesses viewing a perpetrator-
present lineup with those of incorrect witnesses viewing a perpetrator-
absent one.

Subjects

Subjects were 131 Cornell University students enrolled in communica-
tions and human development courses. They received extra credit to-
ward their course grades for participation.

Of the 75 witnesses who viewed a perpetrator-present lineup, 36
(48%) made positive identifications of a specific photograph. Of these,
27 (75%) selected the correct photograph. The remaining 9 witnesses
(25%) mistakenly identified another photo. Of the 56 witnesses viewing
a perpetrator-absent lineup, 31 (55%) incorrectly made a positive
identification.

Interestingly, assertions that the perpetrator was absent from the
lineup were slightly more frequent in the perpetrator-present lineup
condition (33 of 75, 44%) than they were in the perpetrator-absent con-
dition (21 of 56, 38%). Of subjects viewing perpetrator-present lineups,
6 (8%) refused to render a decision; of those viewing perpetrator-absent
lineups, 2 (4%) similarly declined to reach a judgment.

Studies 1–4: Results

We examine five separate issues in turn. The first concerns
whether the decision processes reported by witnesses making
accurate positive identifications differed from those described
by those making erroneous positive incriminations. Second, we
focus on the role of response time in discerning eyewitness ac-
curacy. Third, we explore whether supplemental questionnaire
items also discriminated accurate from inaccurate positive
identifications. Fourth, we evaluate alternative explanations for
the differences we found between correct and incorrect identi-
fications. Fifth, we examine witness reactions to perpetrator-
absent lineups. We also tentatively explore whether witnesses
correctly rejecting a perpetrator-absent lineup (i.e., they con-
clude the offender is not present) differ from those incorrectly
rejecting a perpetrator-present lineup.

Decision Processes of Witnesses Making Accurate
Versus Erroneous Positive Identifications

Did witnesses incriminating the correct photograph differ in their
decision processes from those who made a positive identi-
fication of an incorrect photo? To answer this question, we ex-
amined the number of automatic recognition versus process of
elimination responses witnesses endorsed in each study by means
of a 2 (witness accuracy: accurate vs. inaccurate) × 2
(response type: automatic recognition vs. process of elimina-
tion) mixed-model analysis of variance (ANOVA). 1

The analysis revealed that accurate and inaccurate eyewitnesses
tended to articulate different decision processes, as indi-
cated by a Witness Accuracy × Response Type interaction that
reached statistical significance in each of the four investigations,
F(1, 244) = 6.00, ps < .01.² As shown in Table 1, across the four stud-
ies, accurate witnesses tended to endorse a greater number of
automatic recognition responses (M = .64) than did inaccurate
witnesses (M = .25), t(244) = 4.43, p < .0001.³ As can also be
seen in Table 1, this trend was significant in three of the four
studies, except for Study 1, in which the trend achieved mar-
ginal statistical significance. In contrast, erroneous witnesses
across the four studies more frequently endorsed process of
elimination responses than did their accurate counterparts (Ms
= 2.09 and 1.38 for inaccurate and accurate witnesses, respec-
tively), t(244) = 5.11, p < .0001, and did so significantly in each
individual study.

Collapsing the data across studies, significant differences be-
tween accurate and inaccurate subject-witnesses were also evi-
dent at the individual item level.⁴ As can be seen in Table 2,
accurate and inaccurate witnesses significantly differed from
each other on all five individual items used to assess the auto-
matic recognition versus process of elimination distinction. Accu-
curate witnesses were more likely than inaccurate ones to state
that they “just recognized him” and could not explain why (p < .01), and they were more likely to assert that the perpetrator’s
face had just “popped out” at them (p < .005). Erroneous wit-
nesses were more likely than accurate witnesses to state that
they had compared the photos to each other to narrow the
choices (p < .01), that they had first eliminated some photos
before choosing among the rest (p < .005), and that they had
chosen someone who was close but not exact to their memory
of the perpetrator (p < .0005).

As can also be seen in Table 2, accurate and inaccurate wit-
nesses did not differ in their reports concerning trying to match

---

1 Two notes are necessary about these ANOVAs. First, some readers
may have reservations about applying ANOVAs to measures that can
take on only a few values (e.g., the number of automatic recognition
responses each subject can give is only 0, 1, or 2), thus violating
the assumption of normality. According to conventional statistical wisdom
(e.g., Winer, 1971), this concern is appropriate for studies with a small
n, but constrained scales do not present a significant problem with a
large n, due to the efficiency of the central limit theorem. Second, to
confirm that the first five responses adequately reflected an automatic
recognition versus process of elimination dimension, subjects’ re-
sponses in Studies 1–3 were submitted to a principal-components anal-
ysis. The rotated solution of the analysis revealed that the first five items
loaded together on a single factor, with the process of elimination items
loading positively (loadings > .63) and the automatic recognition re-
sponses loading negatively (loadings < −.62). The remaining two re-
sponses loaded onto a second factor in this analysis. (The “other” option
was not included in this analysis because it was rarely endorsed by sub-
jects across the four studies.)

2 Analyses in the text include all inaccurate subjects participating in
Study 4, those making erroneous positive identifications from perpetra-
tor-absent lineups as well as those reaching incorrect incriminations
from perpetrator-present lineups, unless explicitly noted. When sub-
jects confronting perpetrator-present lineups are excluded, the same
results are obtained in Study 4.

3 All statistical tests reported in this article are two-tailed, except
where otherwise expressly noted.

4 When individual study is included as a factor in all statistical analy-
ses that combine the data across the four studies, no results reported in
the text are altered.
greater number of process of elimination items took longer to accuracy reached statistical significance in Study 3, and was tended to take less time to make their identifications (raw $M = 38.4$ s) than did inaccurate ones (raw $M = 63.9$ s), $t(137) = -3.07$, $p < .005$. This relationship between response time and accuracy reached statistical significance in Study 3, and was supported by clear but nonsignificant trends in Studies 2 and 4. This difference, however, was accounted for by the decision process reported by subject-witnesses. When we control for the number of automatic recognition and process of elimination responses that participants provided, we found that the relationship between response time and accuracy becomes nonsignificant, $t(137) = -.35$, $ns$.

**Response Time**

We explored witness response time, measured in Studies 2–4, to address two separate issues. First, we used response time to assess the validity of our measures of decision processes. To the extent that a witness reported experiencing automatic recognition processes, he or she should make an identification more quickly. To the degree that the witness reported process of elimination strategies, he or she should make an identification more slowly. Second, we explored any potential relationship between response time and accuracy. Do accurate eyewitnesses reach their judgments more quickly than do their erroneous counterparts? Because response times tended to be positively skewed, we conducted all analyses on logarithmic transforms.

Subjects' reports about their decision processes were significantly related to response time, providing some evidence about the validity of our measures. As can be seen in Table 3, over the three studies that included response time measures, subjects endorsing a greater number of automatic recognition items were also more likely to make their identifications quickly, $r(139) = -.52$, $p < .0005$. In addition, subjects endorsing a greater number of process of elimination items took longer to reach their identification decisions, $r(139) = .65, p < .0005$.

Given these findings, it is not surprising to find a significant relationship between response time and eyewitness accuracy (see Table 4). That is, across Studies 2–4, accurate witnesses tended to take less time to make their identifications (raw $M = 38.4$ s) than did inaccurate ones (raw $M = 63.9$ s), $t(137) = -3.07$, $p < .005$. This relationship between response time and accuracy reached statistical significance in Study 3, and was supported by clear but nonsignificant trends in Studies 2 and 4. This difference, however, was accounted for by the decision process reported by subject-witnesses. When we control for the number of automatic recognition and process of elimination responses that participants provided, we found that the relationship between response time and accuracy becomes nonsignificant, $t(137) = -.35$, $ns$.

**Differences on Supplemental Measures**

The third issue of interest concerned whether accurate and inaccurate subjects differed on any of the other responses contained in the post-identification questionnaire. Combining the data from the four studies provided the most sensitive test possible concerning which of these differences could be considered reliable. As can be seen in Table 5, some differences between accurate and inaccurate positive identifications become statistically reliable when combining the data of the different studies, though we should note that these findings rarely achieved statistical significance within any individual investigation.

In particular, accurate and inaccurate subject-witnesses differed in their reports about the impact of nonchosen photos. As shown in Table 5, accurate witnesses across the four studies were more likely to state that nonchosen photos had “little influence” on their decisions ($p < .05$) than were erroneous subjects. Inaccurate subjects, on the other hand, were more likely to assert that nonchosen photographs had “confused” them and “made the task more difficult” ($p < .05$). To a marginal degree,

---

**Table 1**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Witness type</th>
<th>t</th>
<th>F interaction*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accurate</td>
<td>Inaccurate</td>
<td></td>
</tr>
<tr>
<td><strong>Automatic recognition</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 1</td>
<td>.66</td>
<td>.36</td>
<td>1.77****</td>
</tr>
<tr>
<td>Study 2</td>
<td>.62</td>
<td>.13</td>
<td>2.20**</td>
</tr>
<tr>
<td>Study 3</td>
<td>.62</td>
<td>.18</td>
<td>1.32**</td>
</tr>
<tr>
<td>Study 4</td>
<td>.63</td>
<td>.25</td>
<td>2.52**</td>
</tr>
<tr>
<td>Combinedb</td>
<td>.64</td>
<td>.25</td>
<td>4.43***</td>
</tr>
<tr>
<td><strong>Process of elimination</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 1</td>
<td>1.37</td>
<td>2.04</td>
<td>-2.69**</td>
</tr>
<tr>
<td>Study 2</td>
<td>1.46</td>
<td>2.33</td>
<td>-2.33**</td>
</tr>
<tr>
<td>Study 3</td>
<td>1.29</td>
<td>1.94</td>
<td>-2.08**</td>
</tr>
<tr>
<td>Study 4</td>
<td>1.44</td>
<td>2.10</td>
<td>-2.53**</td>
</tr>
<tr>
<td>Combinedb</td>
<td>1.38</td>
<td>2.09</td>
<td>-5.11******</td>
</tr>
</tbody>
</table>

* This statistic refers to the witness accuracy $\times$ response type interaction. ** Collapsing data across all studies. There were 146 accurate positive identifications and 100 inaccurate ones.

$p < .10$. **$p < .05$. ***$p < .01$. ****$p < .0001$. 

---

5 Due to mechanical failure and experimenter error, response times were not available for 5 subjects in Study 2 (3 accurate and 2 inaccurate) and for 10 in Study 3 (7 accurate and 3 inaccurate). These response times were assessed by an individual blind to the hypothesis under study. A second individual, also blind, timed 77% of tape-recorded protocols. Interjudge agreement on these response times was .98.
Table 2
Proportion of Accurate and Inaccurate Eyewitnesses Endorsing Individual Items Concerning Decision Processes (Studies 1–4 Combined)

<table>
<thead>
<tr>
<th>Witness type</th>
<th>Measure</th>
<th>Accurate (n = 146)</th>
<th>Inaccurate (n = 100)</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic recognition</td>
<td>I just recognized him, I cannot explain why.</td>
<td>.25</td>
<td>.11</td>
<td>2.78*</td>
</tr>
<tr>
<td></td>
<td>His face just “popped out” at me.</td>
<td>.38</td>
<td>.14</td>
<td>4.16**</td>
</tr>
<tr>
<td>Process of elimination</td>
<td>I compared the photos to each other in order to narrow the choices.</td>
<td>.49</td>
<td>.67</td>
<td>-2.75*</td>
</tr>
<tr>
<td></td>
<td>I first eliminated the ones definitely not him, then chose among the rest.</td>
<td>.54</td>
<td>.74</td>
<td>-3.16**</td>
</tr>
<tr>
<td></td>
<td>He was the closest person to what I remember, but not exact.</td>
<td>.35</td>
<td>.68</td>
<td>-5.10**</td>
</tr>
<tr>
<td>Other measures</td>
<td>I matched the image in my head to the picture in front of me.</td>
<td>.70</td>
<td>.65</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>I based the judgment on specific facial features.</td>
<td>.53</td>
<td>.62</td>
<td>-1.44</td>
</tr>
</tbody>
</table>

*p < .01.  **p < .0005.

Inaccurate subjects were also more likely to state that the photos “were all so similar that they made me less confident” (p < .10, two-tailed).

Accurate and inaccurate witnesses also differed concerning which had a greater impact on their decision: their memories or the photographs. Across the three studies, accurate witnesses were more likely than their inaccurate peers to state that their memories had a greater influence on their decision (p < .0005), whereas inaccurate subjects were more likely than their accurate counterparts to state that the pictures had a greater impact (p < .0005).

Finally, across the four studies, we should note that we observed a surprisingly strong and unexpected relationship between witness confidence and accuracy. Combining the data of all four studies, we found that witnesses making accurate positive identifications expressed more confidence in their judgments (M = 79.2) than did those making inaccurate identifications (M = 60.0), t(244) = 7.24, p < .0005. Indeed, this relationship was observed in each individual study, all ts > 3, ps < .01.

Table 3
Correlation of Response Time to Reports of Automatic Recognition and Process of Elimination

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study 2</th>
<th>Study 3</th>
<th>Study 4</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic recognition</td>
<td>-.44*</td>
<td>-.73**</td>
<td>-.44**</td>
<td>-.52**</td>
</tr>
<tr>
<td>Process of elimination</td>
<td>.62**</td>
<td>.67**</td>
<td>.64**</td>
<td>.65**</td>
</tr>
</tbody>
</table>

Note. Mock witnesses’ response times were subjected to logarithmic transformations before computing correlations.

* p < .05.  ** p < .0005.

Tests of Alternative Accounts

Taken together, the data presented above suggest that accurate and inaccurate eyewitnesses reach their positive identifications by means of differing cognitive avenues. Accurate positive identifications are more likely to be a product of automatic recognition processes, whereas erroneous incriminations are more likely to be prompted by process of elimination strategies. However, there do exist two alternative accounts for the findings reported above.

The first account pertains to the notion of criterion level (see Gonzalez, Ellsworth, & Pembroke, 1993). Perhaps some witnesses, when faced with the eyewitness identification task, set stringent criteria levels for themselves before making positive identifications. That is, they had to be fairly certain that they were selecting the correct photograph before providing an identification. Others may have used more lenient criteria. It is plausible that witnesses who set stringent criteria may have been more likely to be accurate and to make different reports of decision processes than their more lenient peers. Consequently, differences in decision reports offered by accurate and inaccurate eyewitnesses may reflect nothing more than the use of differing criteria levels.

6 Responses to the question concerning the impact of nonchosen photos were subjected to a principal-components analysis. This analysis revealed two differing factors in the rotated solution that failed to fit nicely into the “little” versus “large” influence dimension we originally had in mind. The first factor involved three items: photos “made the task more difficult,” photos were “so similar” (both loaded > .80), and photos “reinforce[d] my decision” (loading = -.59). The second factor consisted of: photos “had little influence” (loading = .90) and photos “reinforce[d] my decision” (loading = -.74). Because the factors observed failed to match our expectations, we changed our treatment of the items. We felt comfortable combining the items that suggested that the photos had an impact (i.e., they made the task difficult, were so similar). However, we treated the no influence items separately, given that they failed to load on a single factor.
Table 4

Response Times of Witnesses Making Accurate Versus Inaccurate Positive Identifications

<table>
<thead>
<tr>
<th>Witness type</th>
<th>Study</th>
<th>Accurate</th>
<th>Inaccurate</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study 2</td>
<td>39.7</td>
<td>46.5</td>
<td>-8.8</td>
</tr>
<tr>
<td></td>
<td>Study 3</td>
<td>32.3</td>
<td>44.9</td>
<td>-2.36*</td>
</tr>
<tr>
<td></td>
<td>Study 4</td>
<td>45.3</td>
<td>77.9</td>
<td>-1.39</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td>38.4</td>
<td>63.9</td>
<td>-3.07**</td>
</tr>
</tbody>
</table>

* Raw means are reported in table. Comparisons (t tests) were performed after reaction times had been subjected to logarithmic transformations.

** p < .05. *** p < .0005.

Study 3 was designed to test this alternative account directly. Subject-witnesses were given either a high or a low criterion to match before making positive identifications. If criterion level is the key factor producing the correlation between reports of decision processes and accuracy, three predictions can be made. First, subjects given a stringent criterion should report a greater degree of automatic recognition processes than those given a lenient standard. Second, lenient-criterion subjects should report more process of elimination strategies than their high-criterion counterparts. Finally, high-criterion subjects should be more likely to make accurate identifications.

Preliminary analyses revealed that the high- and low-criterion subjects in Study 3 did differ on some measures. First, as expected, a lower proportion of high-criterion subjects made positive identifications (58%) than did low-criterion participants (100%), z = -3.82, p < .005. When high-criterion subjects made positive identifications, they did so with greater confidence (M = 83.0) than their low-criterion counterparts (M = 73.9), t(44) = 1.79, p < .04, one-tailed.

However, stringent- and lenient-criterion subjects did not differ in their reports of automatic recognition versus process of elimination experiences. Reports of decision processes were submitted to a 2 (criterion level) X 2 (witness accuracy) ANOVA. High-criterion subjects were more likely than their erroneous peers to report automatic recognition processes, F = .40. Indeed, they were slightly less likely to do so (Ms = .40 and .50 for high- and low-criterion groups, respectively). In addition, high-criterion subjects were not more likely than their low-criterion counterparts to report process of elimination strategies (Ms = 1.46 and 1.58 for high- and low-criterion groups, respectively), F = .03. Given these findings, it is not surprising that although high-criterion subjects were more likely to be accurate (73%) than were low-criterion subjects (58%), this difference failed to reach statistical significance, z = 1.09.

Although criterion level failed to have an impact on decision processes, those processes still remained correlated with accuracy. Accurate witnesses endorsed a greater number of automatic recognition reports than their erroneous peers (Ms = .61 and .18 for accurate and inaccurate witnesses, respectively), F(1, 44) = 5.12, p < .05. In contrast, inaccurate witnesses endorsed a greater number of process of elimination responses than their accurate counterparts (Ms = 1.94 and 1.29 for inaccurate and accurate witnesses, respectively), F(1, 44) = 3.63, p < .07, albeit to a marginal degree in this analysis. The overall interaction between accuracy and reports of decision processes also remained intact, F(1, 44) = 5.34, p < .05.

The second alternative account pertains to witness confidence. To our surprise, witness confidence was strongly related to accuracy across the four studies, r_p(244) = .42, p < .0001. Perhaps reports of decision processes reflected nothing more than the witnesses' general level of confidence toward his or her judgment. That is, because they had better memories of the perpetrator, or more information about the task presented, highly confident witnesses were more likely to be accurate and to avoid using process of elimination. Buttressing this argument is the relationship we observed between confidence and reports of de-

Table 5

Proportion of Accurate and Inaccurate Eyewitnesses Endorsing Supplemental Items Concerning Decision Processes (Studies 1-4 Combined)

<table>
<thead>
<tr>
<th>Witness type</th>
<th>Measure</th>
<th>Accurate (n = 146)</th>
<th>Inaccurate (n = 100)</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Impact of other photos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>They helped me confirm, reinforce my decision after I made it.</td>
<td>.56</td>
<td>.52</td>
<td>.54</td>
</tr>
<tr>
<td></td>
<td>They had little influence on my decision.</td>
<td>.27</td>
<td>.14</td>
<td>2.38**</td>
</tr>
<tr>
<td></td>
<td>They confused me; they made the task more difficult.</td>
<td>.21</td>
<td>.36</td>
<td>-2.56**</td>
</tr>
<tr>
<td></td>
<td>They were all so similar that they made me less confident.</td>
<td>.19</td>
<td>.29</td>
<td>-1.93*</td>
</tr>
<tr>
<td></td>
<td>Which had a greater influence on the decision</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Memory</td>
<td>.64</td>
<td>.43</td>
<td>3.37***</td>
</tr>
<tr>
<td></td>
<td>The pictures</td>
<td>.08</td>
<td>.26</td>
<td>-3.98***</td>
</tr>
<tr>
<td></td>
<td>They had about equal impact</td>
<td>.27</td>
<td>.32</td>
<td>-7.8</td>
</tr>
</tbody>
</table>

* p < .10. ** p < .05. *** p < .0005.
cision processes. Endorsements of automatic recognition were positively correlated with confidence, \( r(244) = .37, p < .0001 \), and reports of process of elimination items were negatively correlated, \( r(244) = -.47, p < .0001 \).

This account, however, fails to provide a complete explanation for the relationship between witness accuracy and descriptions of decision processes. This fact becomes clear when we statistically control for the relationship of confidence with decision processes before comparing accurate witnesses to inaccurate ones. When we control for confidence, the overall interaction (based on the combined data across the four studies) between witness accuracy and reports of decision processes still holds, \( F(1, 243) = 6.66, p < .02 \). Controlling for confidence, accurate witnesses still report a greater number of automatic recognition processes than their inaccurate peers, \( t(243) = 2.12, p < .05 \). Inaccurate subjects still endorse a greater number of process of elimination responses than do accurate participants, \( t(243) = 2.12, p < .05 \). In short, reports of decision processes still lent independently predictability to eyewitness accuracy beyond that furnished by witness confidence.

**Reactions to Perpetrator-Present Versus Perpetrator-Absent Lineups**

Study 4 examined one last factor that may influence the relationship we found between accuracy and decision processes. In that study, some subjects were never given an opportunity to provide an accurate positive identification because they confronted perpetrator-absent lineups. To what extent did subject-witnesses making false positive judgments from a perpetrator-absent lineup report decision processes like inaccurate subjects who had viewed a perpetrator-present one? And to what degree would they provide decision reports that diverged from eyewitnesses making accurate positive identifications?

In Study 4, when we combine the two inaccurate groups (those making an error upon seeing a perpetrator-present lineup and those reaching erroneous judgments after viewing a perpetrator-absent one), we find the same overall difference in decision reports between accurate and inaccurate eyewitnesses. That is, a 2 (witness type: accurate vs. inaccurate) x 2 (response type: automatic recognition vs. process of elimination) mixed-model ANOVA again revealed the predicted interaction between witness accuracy and the decision process articulated, \( F(1, 65) = 8.69, p < .005 \).

These differences were still apparent when we compared the responses of accurate witnesses to only those making erroneous identifications in the perpetrator-present lineup condition, interaction \( F(1, 34) = 6.54, p < .02 \). As is displayed in Table 6, accurate witnesses were more likely to endorse automatic recognition responses, \( t(34) = 2.16, p < .05 \), and were marginally less likely to endorse process of elimination items, \( t(34) = -1.96, p < .06 \), than were inaccurate subjects. Accurate subjects also differed from their peers who made incorrect positive identifications from perpetrator-absent lineups, interaction \( F(1, 56) = 6.12, p < .02 \). Accurate subjects endorsed a greater number of automatic recognition responses than did these inaccurate participants, \( t(56) = 2.02, p < .05 \), and also endorsed fewer process of elimination items, \( t(56) = -2.16, p < .05 \). Inaccurate subjects viewing perpetrator-present and perpetrator-absent lineups did not differ in their decision reports, \( F < 1 \).

These differences were also evident when we examined the judgments of most interest to forensic officials, identifications of the suspect (of the perpetrator in the perpetrator-present lineup or of his replacement in the perpetrator-absent one). Ten witnesses in the perpetrator-absent condition mistakenly identified the perpetrator's substitute. The decision processes that these subjects described differed significantly from those reported by participants who correctly identified the perpetrator in the perpetrator-present lineup, interaction \( F(1, 35) = 5.43 \). To a marginal degree, these 10 witnesses endorsed a greater number of process of elimination responses \( (M = 2.20), t(35) = 1.95, p < .06 \), and a lesser number of automatic recognition items \( (M = .20), t(35) = -1.84, p < .08 \). However, one caveat is in order about this comparison. These differences fail to appear statistically when confidence is controlled for, interaction \( F(1, 34) = 1.64 \).

Study 4 also allowed us to examine whether witnesses making accurate and inaccurate nonidentifications also differed in their decision processes, albeit in an exploratory manner. Specifically, on a number of measures, we compared the responses of subjects who accurately stated that the perpetrator was not present in the perpetrator-absent lineup with those who inaccurately rejected lineups containing the perpetrator. Regrettably, we found few differences between these sets of subjects. Although witnesses making accurate nonidentifications reached their decisions more quickly \( (raw M = 61.9 \text{ s}) \) than their inaccurate peers \( (raw M = 67.1 \text{ s}) \), this difference failed to reach statistical significance, \( t = -1.03, ns \), after logarithmic transforms. In terms of confidence, witnesses making accurate nonidentifications were not any more or less confident \( (M = 71.8) \) than their inaccurate counterparts \( (M = 75.0), t < 1 \).

When we examine the responses of witnesses making nonidentifications on our close-ended queries about decision processes, we again find few differences. This is not surprising, given that many of these questions were aimed at exploring the decision processes experienced by participants making positive identifications. However, there are two differences of note, both regarding the question probing what had had a greater impact on witnesses' decisions: their memories of the perpetrator or the

<table>
<thead>
<tr>
<th>Measure</th>
<th>Accurate (n = 27)</th>
<th>Offender present (n = 9)</th>
<th>Offender absent (n = 31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic recognition</td>
<td>.63</td>
<td>.11</td>
<td>.29</td>
</tr>
<tr>
<td>Process of elimination</td>
<td>1.44</td>
<td>2.22</td>
<td>2.07</td>
</tr>
</tbody>
</table>

---

**Table 6 Mean Number of Automatic Recognition and Process of Elimination Responses Endorsed by Accurate and Inaccurate Eyewitnesses Viewing Perpetrator-Present and Perpetrator-Absent Lineups in Study 4**
photographs presented. Accurate witnesses were less likely to state that their memories had had a greater impact (57%) than were inaccurate ones (82%), $z = -2.06, p < .05$. Instead, to a marginal degree, accurate witnesses more frequently stated that their memories and the pictures had an equal impact on their decisions (35% vs. 15% for accurate and inaccurate subjects, respectively), $z = 1.71, p < .09$.

### Summary

In sum, in all four studies, eyewitnesses making accurate identifications of the perpetrator could be distinguished in part from witnesses making erroneous positive identifications by asking them to describe their decision processes. Accurate witnesses were more likely to assert, relative to inaccurate peers, that their decisions were a product of automatic recognition. Inaccurate witnesses, when compared with accurate counterparts, were more likely to state that their decisions were based on process of elimination strategies.

Supplemental analyses also provided evidence consistent with those differences. Accurate witnesses reached their decisions more quickly than their inaccurate peers. They were also marginally more likely than their inaccurate counterparts to say that nonchosen photos had little influence on their judgments, whereas inaccurate eyewitnesses were more likely to state that the pictures contained in the lineup had confused them or made their decisions more difficult. Accurate witnesses were also more likely than their inaccurate peers to assert that their memories had had a greater impact on their decisions than the pictures did. Inaccurate witnesses were more likely to claim the opposite.

### Study 5

How useful is the information garnered in the previous four studies? If subjects were placed in the role of police officer, lawyer, judge, or juror, and asked to decide whether a given eyewitness identification was accurate or erroneous, would they find cues about decision processes to be helpful? Would knowledge of a witness’s decision processes prompt people to distinguish accurate from inaccurate eyewitnesses more successfully?

Study 5 was designed to address these issues. Subjects were given a series of post-identification questionnaires completed by some of the subject-witnesses (in Studies 1, 2, and 4) and were asked, for each, to decide whether the witness’s positive identification had been correct or incorrect. In this way, we could discern whether people intuitively and successfully use knowledge of witness decision processes when assessing witness accuracy.

In addition, Study 5 explored whether giving subjects explicit knowledge of the automatic recognition versus process of elimination distinction would enhance their performance. Roughly half the subjects in Study 5 were informed of the findings of the previous three studies, being told that accurate witnesses tended to reach their identifications by means of automatic recognition and that inaccurate witnesses tended to pursue process of elimination strategies. We then compared the success rates of this informed group to those achieved by uninformed subjects.

### Method

#### Subjects

Subjects were 26 students attending Cornell University summer session courses in psychology. Subjects either received extra credit toward their course grades or were paid $4.

#### Procedure

Upon arriving at the lab, subjects were instructed that the study focused on the psychology of eyewitness testimony. They were given a brief description of the crime and identification task used in Studies 1, 2, or 4, and were told that they would be looking at questionnaires completed by 40 different witnesses who had made positive identifications. They were told that their task was to decide which of the witnesses had made accurate identifications and which had rendered erroneous ones. Subjects were then handed a folder that contained a consent form, further instructions, the 40 witness protocols, and a sheet for marking their judgments.

The folders contained two different sets of instructions. In the informed condition ($n = 12$), the instructions included important hints to consider when making judgments. Informed subjects were told that accurate witnesses tended to endorse automatic recognition responses and to state that their memories had had more impact on their decisions than the lineup pictures. Informed participants were further instructed that inaccurate witnesses tended to cite process of elimination items and to report that the pictures had a greater influence on their decision than their memories. These specific instructions were selected because they represented differences between accurate and inaccurate witnesses that were clearly evident in the previous three studies (see Table 5). Informed subjects were further told that these hints were to be considered “helpful, but not to be compelling,” and that they “should use [their] own best judgment in addition to these hints” when rendering their decisions. In the uninformed condition ($n = 14$), the instructions contained none of these hints; they merely reiterated that subjects were to decide whether each witness had been accurate or inaccurate.

After making judgments for the 40 witness protocols provided, subjects answered an open-ended question asking them to describe the strategies they followed when completing the task. They also completed a questionnaire asking them which components of witness testimony they had found most helpful. Specifically, they were presented with all 14 of the close-ended items contained in the post-identification questionnaire. For each item, they were asked whether a witness endorsing it was more likely to be accurate or inaccurate in his or her identification.

Subjects responded on an 11-point scale that ranged from $-5$ (witnesses endorsing this item were definitely inaccurate) to 0 (this item has nothing to do with accuracy or error) to 5 (witnesses endorsing this item were definitely accurate).

After completing these questionnaires, subjects were debriefed concerning the rationale and aims of the study. Interested subjects were given answer sheets for the task.

#### Materials

Each subject was given information from questionnaires completed by 40 subject-witnesses making identifications in Studies 1, 2, and 4. There were two different packets of 40 protocols, each containing questionnaires from 20 accurate and 20 inaccurate eyewitnesses. The specific accurate and inaccurate protocols included in each packet were selected at random from all subjects making identifications in three of...
the previous studies. Once selected, the protocols in each replication set were placed in a random order.

Subjects in Study 5 saw only a portion of the information provided by each eyewitness. Specifically, they were given information concerning how each witness had responded to the three close-ended questions included in the previous studies ("How would you describe decision making?" "How much influence did the other pictures have?" and "Did memory or pictures influence decision more?"). The specific photograph chosen, witness confidence, and responses to open-ended queries were not included because we wished to see whether subjects would rely on the specific measures we had found successfully differentiated accurate and inaccurate identifications.

Results and Discussion

Subjects displayed some ability to differentiate accurate from inaccurate eyewitness identifications. As can be seen in Table 7, uninformed subjects correctly classified 60.7% of witness protocols, a rate well above chance levels, \(t(13) = 5.34, p < .0005\). However, this rate of success was significantly below the performance levels attained in the informed condition, \(M = 65.8\), as evidenced by a main effect for instruction, \(F(1, 22) = 4.76, p < .05\), in a 2 (instruction: informed vs. uninformed) \(\times 2\) (replication set: A vs. B) ANOVA. Inspection of Table 7, however, reveals one caveat about this finding. Informed subjects outperformed uninformed participants (by 10%) mostly when they judged protocols coming from witnesses making inaccurate identifications, \(F(1, 22) = 3.37, p < .05\). Both groups of subjects achieved nearly identical performance levels (roughly 64%) when making decisions about accurate witness protocols. Subject accuracy did not differ across the two replication sets.

How did subjects reach these levels of accuracy in their decisions, and how did the informed group's instructions prompt them to attain significantly higher levels of performance? To determine this, we conducted further analyses that focused on the factors subjects gave weight to in their decisions. For each subject, we regressed his or her 40 decisions on six measures: (a) number of automatic recognition items cited in each protocol, (b) number of process of elimination responses made, (c) number of responses stating that the nonchosen photos had had negligible impact, (d) number of responses stating that the nonchosen photos had made the task more difficult, (e) whether the witness stated that memory had more impact than the photos, and (f) whether the witness stated that the photos were more influential than memory.

Table 7

<table>
<thead>
<tr>
<th>Accuracy of witness</th>
<th>Condition</th>
<th>Informed ((n = 12))</th>
<th>Uninformed ((n = 14))</th>
<th>(F(1, 22))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td></td>
<td>64.2</td>
<td>63.8</td>
<td>.00</td>
</tr>
<tr>
<td>Inaccurate</td>
<td></td>
<td>67.5</td>
<td>57.5</td>
<td>3.37**</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>65.8</td>
<td>60.7</td>
<td>4.76**</td>
</tr>
</tbody>
</table>

* \(p < .05\). ** \(p < .05\).

This analysis revealed that uninformed subjects possessed a partial understanding of the cognitive processes that differentiated accurate from inaccurate witnesses (see top panel of Table 8). Specifically, subjects in the uninformed condition did give weight to whether witnesses cited automatic recognition responses, albeit to a marginal degree. That is, they were more likely to judge a witness accurate to the extent that the witness endorsed automatic recognition responses, mean \(\beta = .08, t(13) = 1.95, p < .05\). However, uninformed subjects did not give sig-

Another gauge of subject performance is to compare their performances to an accuracy rate achieved by a computer "trained" to take note of the responses we mentioned in the instructions of the informed condition. Toward that end, we computed an accuracy index for each of the 80 witness protocols. This index consisted of the number of accuracy-related responses the witness endorsed (3) minus the number of accuracy-related items endorsed (4). Witnesses were judged to be accurate when their accuracy index scores were at or above the median (−1). They were judged to be inaccurate when their index scores lay below the median. When this procedure was used, 53 of the 80 witnesses (66.3%) were correctly classified (63.8% of accurate witnesses and 69.7% of inaccurate ones). Comparing subjects' success rates in the informed condition to these figures reveals that these subjects were performing at near optimum levels.

<table>
<thead>
<tr>
<th>Questionnaire measure</th>
<th>Condition</th>
<th>Informed</th>
<th>Uninformed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual regression weights ((b)) in judges' accuracy assessments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatic recognition</td>
<td>(.21***)</td>
<td>(.08)</td>
<td></td>
</tr>
<tr>
<td>Process of elimination</td>
<td>(-.25***)</td>
<td>(-.08)</td>
<td></td>
</tr>
<tr>
<td>Nonchosen photos have negligible impact</td>
<td>(.10)</td>
<td>(.20***)</td>
<td></td>
</tr>
<tr>
<td>Nonchosen photos made task more difficult</td>
<td>(-.25***)</td>
<td>(-.36***)</td>
<td></td>
</tr>
<tr>
<td>Memory had more impact than photos</td>
<td>(.30***)</td>
<td>(.09**)</td>
<td></td>
</tr>
<tr>
<td>Photos had more impact than memory</td>
<td>(-.01)</td>
<td>(-.05)</td>
<td></td>
</tr>
</tbody>
</table>

* For judges' stated beliefs, responses were made on a scale ranging from \(-5\) (witness definitely inaccurate when endorsed) to 0 (response has no relationship to accuracy) to 5 (witness was definitely accurate). ** \(p < .05\). *** \(p < .005\).

8 Another gauge of subject performance is to compare their performances to an accuracy rate achieved by a computer "trained" to take note of the responses we mentioned in the instructions of the informed condition. Toward that end, we computed an accuracy index for each of the 80 witness protocols. This index consisted of the number of accuracy-related responses the witness endorsed (3) minus the number of accuracy-related items endorsed (4). Witnesses were judged to be accurate when their accuracy index scores were at or above the median (−1). They were judged to be inaccurate when their index scores lay below the median. When this procedure was used, 53 of the 80 witnesses (66.3%) were correctly classified (63.8% of accurate witnesses and 69.7% of inaccurate ones). Comparing subjects' success rates in the informed condition to these figures reveals that these subjects were performing at near optimum levels.
significant weight to process of elimination responses, mean $\beta =-.08, t = -1.51, ns.$ In contrast, informed witnesses tended to give greater weight to both of these factors. They were significantly more likely to judge witnesses accurate when they cited automatic recognition responses, mean $\beta =.21, t(11) = 8.18, p < .0005.$ They were significantly less likely to label witnesses as accurate when they endorsed process of elimination items, mean $\beta =-.25, t(11) = 5.36, p < .005$.

Regarding the query on which item had more influence, uninformed subjects gave significant weight to whether witnesses stated that their memories had more impact than the pictures provided, mean $\beta = .09, t(13) = 2.25, p < .05,$ but not to whether witnesses said that the pictures had had more impact than their memories, mean $\beta =.05, t = -1.15, ns.$ Informed witnesses acted similarly. When witnesses stated that their memories had been more influential than the pictures, informed subjects were more likely to judge them to be accurate, mean $\beta = .30, t(11) = 8.52, p < .0005.$ Informed subjects, however, did not give weight to when witnesses said that the pictures had more impact, mean $\beta = -0.01, t < 1$.

One final note is necessary regarding these regression analyses. They revealed that informed subjects gave all of these factors mentioned in their instructions, except the “pictures had more impact than memory” response, significantly more weight than did their uninformed peers: $F(1, 22) = 9.00, p < .01,$ for automatic recognition; $F(1, 22) = 5.50, p < .03,$ for process of elimination; and $F(1, 22) = 17.84, p < .005,$ for memory has more impact.

Were subjects aware of the factors they were giving weight to in their decisions? Did they possess explicit knowledge of the factors that could help discern accurate from inaccurate eyewitnesses? Recall that we asked subjects about their theories concerning the relation of cognitive processes to eyewitness accuracy. From these measures, we derived assessments of subjects’ beliefs about the six factors examined in the aforementioned regression analysis.

Analyses of these explicit belief measures mirrored the results of the regression analysis (see bottom portion of Table 8). On average, uninformed subjects stated that witnesses citing automatic recognition responses were more likely to be accurate than inaccurate, $M = 2.7,$ a rating significantly different from the no difference midpoint of the scale, $t(13) = 6.29, p < .0005.$ However, they failed to associate process of elimination strategies with inaccuracy. When asked about witnesses endorsing process of elimination responses, uninformed subjects stated to a nonsignificant degree that witnesses making such responses were more likely to be more accurate than less, $M = 0.4, t < 1.$ In contrast, informed subjects articulated beliefs more consistent with the findings of the previous three studies. They stated that witnesses citing automatic recognition items were more likely to be accurate than less, $M = 3.9, t(10) = 15.84, p < .0005,$ and, unlike their uninformed counterparts, asserted that witnesses citing process of elimination responses were less likely to be incorrect than accurate, $M = -1.4, t(10) = -2.32, p < .05.$

In addition, uninformed subjects on average possessed only a partial understanding of the relevance of the question regarding which item had more influence. They stated that witnesses were more likely to be correct when these witnesses asserted that their memories had a greater impact on their identifications than the photos, $M = 3.6, t(13) = 9.79, p < .0005.$ However, they articulated no belief that witnesses citing the photos as more influential would tend to be inaccurate, $M = 0.2, t < 1.$ In contrast, informed witnesses related both measures to witness accuracy. They stated that witnesses citing their own memories as more influential were more likely to be accurate, $M = 3.7, t(10) = 9.72, p < .0005,$ and professed marginally significant beliefs that witnesses stating that the photos had more impact would be less accurate, $M = -1.9, t(10) = -2.08, p < .08.$

In sum, Study 5 provided evidence that people possess a partial intuitive understanding of the differences between accurate and inaccurate eyewitness identifications revealed in the previous three studies. Subjects given no explicit instruction tended to believe that automatic recognition, as well as reliance on memory as opposed to lineup photos, was related to accuracy. They did not associate process of elimination, and reliance on lineup photos, with inaccuracy. As a consequence, they achieved modest, though clear, success at discriminating accurate from inaccurate eyewitness identifications. Giving subjects explicit information about the relationship between decision processes and eyewitness accuracy significantly improved their level of performance. Subjects given such information gave weight to process of elimination responses and more closely correlated their judgments of accuracy with the decision process indicators articulated in the previous four studies.

We should note, however, that Study 5 probably underestimates the ability of people to discriminate accurate from inaccurate eyewitness identifications, given the information we had in hand. To examine rigorously whether subjects understood and could use information about decision processes, we withheld other information that could have helped them achieve success. For example, we withheld information about eyewitness confidence, which in these four studies was strongly related to accuracy. Past research suggests that subjects would have relied on this type of information (Brigham & Bothwell, 1983; Cutler et al., 1988; Kassin et al., 1991; Leippe et al., 1992; Lindsay et al., 1989; Wells et al., 1979). In this instance, information about confidence would have allowed subjects to achieve a higher level of accuracy in their decisions.

**General Discussion**

Eyewitnesses making accurate identifications of the perpetrator from a lineup can be distinguished, in part, from witnesses making erroneous positive identifications by asking them to describe the cognitive processes they experienced while reaching their judgments. In the first four studies, accurate witnesses were more likely to assert, relative to inaccurate peers, that their decisions were a product of automatic recognition. That is, they were relatively unable to describe any explicit cognitive strategy that had led them to the decision they reached. Inaccurate witnesses, in contrast, were more likely to state that their decisions were...
were based on a deliberative process of elimination strategies. They had compared the lineup photographs to each other, eliminating the worst candidates for identification and selecting from the rest.

Supplemental analyses also suggested that accurate positive identifications were accompanied by less explicit use of cognitive operations than were erroneous ones. Replicating Sporer (1993), we found that witnesses making accurate positive identifications reached their judgments more quickly than their peers who made erroneous positive identifications, although reports of decision processes tended to be a more sensitive indicator of accuracy. Accurate witnesses were also more likely to say that nonchosen photographs had little influence on their judgments, whereas inaccurate eyewitnesses were more likely to state that the other lineup pictures had confused them or made their decisions more difficult. Accurate witnesses were also more likely than their inaccurate peers to assert that their memories had more impact on their decisions than did the pictures. Inaccurate witnesses more frequently claimed the opposite.

Along the way, individual studies ruled out alternative accounts for these differences between correct and incorrect eyewitness judgments. Study 3 showed that they were not a result of different response criteria. Requiring subjects to meet stringent or lenient criteria before making their identifications did not prompt them either toward relying on automatic recognition or process of elimination strategies. Moreover, asking subjects to adopt either stringent or lenient judgment criteria did not influence the link between the cognitive operations and identification accuracy. These differences were also not explained by the general level of confidence that subjects expressed in their decisions, despite the fact that accurate witnesses were more confident in their decisions than were inaccurate ones, and even though confidence was positively correlated with automatic recognition responses and negatively associated with reports of process of elimination. When we statistically controlled for the relationship of confidence to accuracy, we still found that witness accuracy was significantly correlated with the cognitive operations that subjects reported.

Finally, data across the first four studies also served to rule out a final alternative account for the differences observed between accurate and inaccurate positive identifications. Specifically, it might be argued that the judgments of accurate witnesses in these studies was a product of an absolute judgment strategy as opposed to automatic recognition. That is, accurate witnesses conducted a deliberate "comparison of each lineup member to one's recollection" of the perpetrator (Lindsay & Wells, 1985, p. 558), without comparing the lineup options to each other. They then positively identified a photograph if that choice was a good match to a witness's memory. Several points in the data argue against this interpretation. First, subjects had the opportunity to endorse a response that is very similar to a description of an absolute judgment strategy (e.g., "I matched the image in my head to the pictures in front of me"), and accurate and inaccurate witnesses did not diverge in the frequency with which they endorsed this response. Second, several assumptions or explanations would have to be added to explain why an absolute judgment strategy would lead subjects to protest that they could not describe their decision-making process ("I just recognize him; I cannot explain why") or describe it as a decision that "just happened" to them ("His face just popped out at me"). In addition, an absolute judgment explanation would not necessarily predict that accurate positive identifications would occur more rapidly than erroneous ones, whereas an automatic recognition account would make that prediction.

One issue should be addressed about our reliance on subjects' reports of their decision processes. Although useful in distinguishing eyewitness accuracy from error, were subjects' reports of their decision processes accurate? Such a concern is appropriate, given many demonstrations that people often are not aware of the factors or cognitive processes that prompt their decisions and actions (Nisbett & Wilson, 1977). We should note, in response, that supplemental data suggest that subjects' reports were accurate. For example, subjects reporting automatic recognition processes reached their identifications relatively quickly. In addition, we deliberately implemented procedures to ensure as much accuracy as possible in subjects' reports. Past work has shown that people are inaccurate in reporting their decision processes when they work retrospectively, that is, describe their decisions after they have been reached. In the preceding experiments, we made sure that subjects talked about their decisions concurrently, while they were making them. According to Ericsson and Simon (1980), such concurrent verbal protocols should provide accurate information about subjects' thoughts while they are making their decisions, as long as the procedure is not too intrusive. By asking subjects to provide these protocols, we ensured that they had a salient memory on which to base their post-identification questionnaire responses. By the same procedure, we avoided prompting them to infer or reason about what that process must have been like, a situation that is most likely to lead to erroneous reports of cognitive processes (Ericsson & Simon, 1980; White, 1980).

Implications for Reality Monitoring

These findings carry many implications for basic research on memory processes, especially work on reality monitoring (Johnson & Raye, 1981). Like past work, we found that accurately reported memories of external events (the correct incrimination of a perpetrator) differed from internally generated memories (the "imagined," mistaken belief that the witness had seen an individual's face before) in terms of the cognitive operations that subjects reported. Accurate recognitions occurred with less deliberative thought than did inaccurate identifications. Like past work (Schooler et al., 1986), subject-witnesses could articulate these differences when asked, and these differences could be used by observers to distinguish accuracy from error.

These studies also extend work on reality monitoring. The work of Schooler et al. (1986) focused on people's descriptions of past memories. As such, it gave little guidance as to how people could distinguish accurate from inaccurate recognition judgments, in which there is no memory to describe, only an object to identify as previously encountered. The results of Studies 1-4 suggest that reality monitoring principles can be used to assess the veracity of recognition judgments. What one
must do is to ask subjects about "the mental road" they traveled when reaching their recognitions. External recognitions (the subject has previously seen a stimulus) and internally generated recognitions (the subject erroneously identifies a stimulus as seen previously) differ on whether subjects cite a number of deliberative cognitive operations in their judgment.

With this fact in mind, it is informative to consider various models of recognition judgment. Several theorists agree that recognition memory involves two different processes. Recognition can be based on general nonanalytic familiarity with an object, a process in which the object is remembered as a "coherent and unitary perceptual event" (Mandler, 1980, p. 255). Or, the object can be recognized after a deliberative, analytic retrieval process in which the individual attempts to recall specific details about the object from memory (Mandler, 1980; see Bartlett, Till, & Levy, 1980; Jacoby, Kelley, & Dywan, 1989, for similar distinctions). What the present studies suggest, at least in the realm of facial recognition, is that one process indicates accuracy in recognition (the general familiarity one), whereas the second (the analytical retrieval process) suggests error. Future research could investigate this possibility, to explore whether relying on a general sense of familiarity is a process or strategy that is related to more accurate judgment in recognition memory.

Witness Confidence

Some surprising aspects of subjects' responses deserve discussion. In Studies 1-4, witness confidence was strongly related to identification accuracy, although it is generally accepted that eyewitness confidence and accuracy are largely uncorrelated (Booth et al., 1987; Smith et al., 1989; Wells & Murray, 1984).

There are several explanations for this unexpected finding, and several reasons to believe that the present experiments overstate the relationship between confidence and accuracy. For example, confidence is more likely to be related to accuracy when examining only positive identifications and ignoring rejections of the lineup (Malpass & Devine, 1981, 1984; Sporer, 1993). In our work we focused primarily on positive identifications from a perpetrator-present lineup, thus concentrating on a circumstance most favorable toward finding a confidence-accuracy link. To assess the relationship between confidence and accuracy more comprehensively, we would have had to focus on many other types of decisions that witnesses could reach (e.g., stating that the perpetrator is not present). Indeed, when we examined non-identifications in Study 4, we found that accurate judgments were made with less confidence than inaccurate ones, albeit a nonsignificant trend.

In addition, the significant correlation between confidence and accuracy may have been prompted by requesting subjects to provide concurrent verbal protocols. That is, asking subjects for think-aloud protocols may have alerted them more strongly to the cognitive operations they experienced when reaching their decisions, and thus provided them with valid cues to consider when assessing confidence. Past work on reality monitoring reveals that people are aware of the qualitative differences between their perceived and imagined memories (Johnson & Raye, 1981). Furthermore, in our studies, confidence was related to the characteristics of decision processes that we found were, indeed, related to accuracy. There is some indirect precedence for these speculations. A similar logic surrounds Kassin's (1985; Kassin et al., 1991) retrospective self-awareness manipulation, in which subjects are asked for confidence estimates only after viewing a videotape of themselves reaching their identification. This procedure has been shown to promote stronger correlations between confidence and accuracy, presumably because it provides subjects with cues associated with accuracy or error.

With this in mind, future work could be done on whether the surprisingly strong correlations we found between confidence and accuracy were prompted by think-aloud procedures. Indeed, this notion could be tested in many domains in which people have demonstrated only a modest ability to anticipate the accuracy or error of their judgments (Dunning, Griffin, Miojkovic, & Ross, 1990; Dunning & Story, 1991; Vallone, Griffin, Lin, & Ross, 1990).

Forensic Implications

These studies also suggest implications for criminal justice officials dealing with eyewitness testimony. They suggest that asking witnesses how they reached their identification judgments may help forensic scientists discern meaningful from meaningless information. This was explicit in Study 5, in which providing observers with information about witness decision processes, and hints concerning what to look for, helped them to differentiate accurate eyewitness identifications from inaccurate ones.

However, three caveats about the forensic application of this research are in order. First, noting whether the witness achieved a positive identification by means of automatic recognition or process of elimination is not a method that works with certainty. Although reports of decision processes helped us and the observer-subjects of Study 5 to make better educated guesses about the veracity of specific eyewitness reports, the procedure did not work 100% of the time. As such, reports of decision processes should be only one of many indicators considered when assessing the accuracy of a specific witness. Second, the results of the four experiments were based on a single set of stimulus materials, using the same crime, thief, and lineup. To ensure the generality of our findings, and their applicability to forensic settings, the link between decision processes and accuracy should be tested with a variety of stimuli and under a range of circumstances. We consistently found such a relationship, but would that link have been stronger or weaker had differing lineups, perpetrators, crimes, and witnesses been used?

Finally, the forensic applicability of the present research depends on one comparison that we were able to conduct only once. That comparison involves witnesses who correctly identify the perpetrator from a perpetrator-present lineup and witnesses who incorrectly identify his or her replacement in a perpetrator-absent one (Lindsay & Wells, 1985). To a police officer dealing with an actual case, identifications of all other lineup choices do not matter. They are known to be innocent foils. In Study 4, we conducted that comparison and found that subjects...
accurately identifying the perpetrator differed from those erroneously incriminating his substitute in a perpetrator-absent lineup, but would that generally be the case? Further work focusing on this specific question is necessary to assess the forensic applicability of the present research.

Other Avenues for Future Research

Other steps could be taken to expand the usefulness of the procedure described in the preceding studies. We concentrated primarily on positive identifications. But what about rejections of the lineup (i.e., “The perpetrator is not there”)? Can accurate rejections be distinguished from incorrect ones? In Study 4, we began to explore this issue, and found few differences between correct and erroneous rejections. Accuracy of rejection was not related to witness confidence, nor with the speed with which subject-witnesses made their decisions.

One difference, however, did appear on post-identification questionnaires. Witnesses incorrectly rejecting a perpetrator-present lineup were more likely to state that their decision was primarily based on their memory of the perpetrator than were participants accurately rejecting a perpetrator-absent lineup. Witnesses correctly rejecting a perpetrator-absent lineup were marginally more likely to state that their decision was equally influenced by their memories and the pictures than were their inaccurate counterparts.

Why might this be so? Perhaps these differences occurred because we examined so many different measures that some, by chance, were bound to appear statistically significant. However, if one assumes that correct rejectors possessed more accurate memory of the perpetrator than did their incorrect counterparts, their differing post-identification responses become explicable. All the lineup alternatives were at least moderately similar to the perpetrator. Thus, when correct rejectors examined the perpetrator-absent lineup, all alternatives were plausible and had to be considered before rejecting the lineup. As a result, correct rejectors would cite both their memories and the pictures as important factors in their decisions. By contrast, if incorrect rejectors had inaccurate memories of the perpetrator, the lineup alternatives would not have been as plausible. As such, they would not match their memory of the perpetrator, and would not be as likely to be cited as an important factor in their decision.

The preceding analysis is speculative, and more attention needs to be paid to differences (if any exist) between accurate and inaccurate rejectors. Suffice it to say that accurate rejection of a lineup may not involve the key ingredient we found in accurate positive identifications, automatic recognition. According to reality monitoring principles, recognition that an object has been previously encountered often occurs because of a “rush” of sensory and perceptual memories. It is difficult to see how the absence of an object would lead to such a rush of a decision.

In addition, future work could focus on witness reactions to other lineup procedures. We did not explore show-ups, in which the police present one individual to the witness and ask, yes or no, whether he or she is the perpetrator. We can predict that accurate positive identifications from show-ups would likely occur from automatic recognition, but what about inaccurate positive identifications? With these judgments, there is no process of elimination strategy to follow, so what explicit cognitive operations might inaccurate witnesses report? And what of sequential lineup procedures? Inaccurate positive identifications do occur with this technique. Might those errors be indicated by some form of explicit cognitive operations? There does exist some tentative evidence that inaccurate positive identifications with sequential lineups involve more deliberate cognitive processes than do accurate ones. Specifically, Sporer (1993) discovered that inaccurate incriminations were made more slowly than accurate ones under the sequential lineup techniques.

Perhaps most important, the findings of these studies suggest a procedure by which the accuracy of eyewitness identifications can be improved. What if we forced witnesses to adopt one decision strategy over another? Would compelling witnesses to adopt an automatic recognition strategy increase the rate of accurate identifications? Would forcing subjects to pursue a process of elimination approach impair their ability to select the correct photograph? Finding procedures to enhance eyewitness accuracy, and learning to avoid ones that increase error, would go a long way toward assisting accurate fact-finding in eyewitness cases. In addition, such procedures, if successful, would sharpen our theoretical understanding of accuracy and error in eyewitness identification. This is an avenue we are currently pursuing.

References


Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The


Received July 13, 1993
Revision received March 14, 1994
Accepted April 12, 1994

---

**P&C Board Appoints Editor for New Journal:**

*Journal of Experimental Psychology: Applied*

In 1995, APA will begin publishing a new journal, the *Journal of Experimental Psychology: Applied*. Raymond S. Nickerson, PhD, has been appointed as editor. Starting immediately, manuscripts should be submitted to

Raymond S. Nickerson, PhD
Editor, *JEP: Applied*
Department of Psychology
Tufts University
Medford, MA 02155

The *Journal of Experimental Psychology: Applied* will publish original empirical investigations in experimental psychology that bridge practically oriented problems and psychological theory. The journal also will publish research aimed at developing and testing of models of cognitive processing or behavior in applied situations, including laboratory and field settings. Review articles will be considered for publication if they contribute significantly to important topics within applied experimental psychology.

Areas of interest include applications of perception, attention, decision making, reasoning, information processing, learning, and performance. Settings may be industrial (such as human–computer interface design), academic (such as intelligent computer-aided instruction), or consumer oriented (such as applications of text comprehension theory to the development or evaluation of product instructions).