Running Head: INVESTIGATING INVESTIGATORS

Investigating Investigators:

Examining the Impact of Eyewitness Identification Evidence on Student-Investigators

Melissa A. Boyce, D. Stephen Lindsay, and C. A. E. Brimacombe

University of Victoria

Melissa A. Boyce, D. Stephen Lindsay, C. A. E. Brimacombe, Department of Psychology, University of Victoria.

Portions of this research were presented at the 2005 and 2006 meetings of the American Psychology-Law Society. This research was supported in part by grants from the Social Sciences and Humanities Research Council of Canada to Melissa A. Boyce, D. Stephen Lindsay and C. A. E. Brimacombe.

Correspondence concerning this article should be addressed to Melissa Boyce, Department of Psychology, University of Victoria, P.O. Box 3050 STN CSC, Victoria, British Columbia, Canada, V8W 3P5; phone: 250-721-6327; fax: 250-472-5014; e-mail: mboyce@uvic.ca.

Keywords: crime investigation; eyewitness identification; investigator decision-making

Abstract

This research examined the impact of eyewitness identification decisions on student-investigators. Undergraduates played the role of police investigators and interviewed student-witnesses who had been shown either a good or poor view of the perpetrator in a video-taped crime. Based on information obtained from the witness, student-investigators then chose a suspect from a database containing information about potential suspects and rated the probability that their suspect was the culprit. Investigators then administered a photo lineup to witnesses, and re-rated the probability that their suspect was guilty. Student-investigators were highly influenced by eyewitness identification decisions, typically overestimating the information gained from the identification decision (except under conditions that led witnesses to be very accurate), and were generally unable to differentiate between accurate and inaccurate witnesses.

Investigating Investigators:

Examining the Impact of Eyewitness Identification Evidence on Student-Investigators

Eyewitnesses often provide the only direct evidence that a person is guilty of committing a crime. Most other evidence is indirect or circumstantial. Indirect evidence, such as fingerprints or DNA, can place a person at the crime scene but cannot always establish that the suspect actually committed the crime. This gives eyewitness testimony a lot of influence in criminal proceedings. In fact, eyewitness testimony has been shown to be among the most compelling forms of evidence to juries (Lindsay, 1994), having as much impact as a confession in some cases (Kassin & Neumann, 1997).

Given that eyewitness evidence plays such a significant role in the justice system, the accuracy of eyewitnesses becomes a very important consideration. Unfortunately, there is a vast amount of research showing that eyewitness evidence can be highly unreliable and inaccurate (Lindsay & Wells, 1985; Wells, 2000). Yet eyewitnesses are not always in error (e.g., Yuille & Cutshall, 1989). This raises the question of how well individuals involved in the criminal justice system can differentiate between accurate and inaccurate eyewitness evidence. In the current research, our interest is focused on how those in the role of investigating officer evaluate eyewitnesses' identification evidence.

Prior research has indicated that mock jurors consistently overestimate eyewitnesses' accuracy and fail to differentiate between accurate and inaccurate eyewitnesses (e.g., Brigham & Bothwell, 1983; Lindsay, Wells, & O'Connor, 1989; Lindsay, Wells, & Rumpel, 1981; Wells, Lindsay, & Tousignant, 1980). There is no research that we know of that has specifically compared investigators to jurors; however, we speculate that investigators may differ from jurors in these regards for several reasons. First, jurors might be biased to assume that an eyewitness is

accurate because he or she is testifying in court; they may presume that the eyewitness's credibility has been verified by the police officers investigating the case, as well as by the prosecution. Furthermore, there may be differences between witnesses who make it to trial and those who don't; jurors see only the former whereas investigators see all. Also, by the time witnesses appear in court, delay and various psychosocial processes may attenuate differences between accurate and inaccurate witnesses (e.g., Wells, Ferguson, & Lindsay, 1981). Finally, jurors passively observe witnesses being examined in the ritualistic routines of the courtroom, which may limit their ability to differentiate accurate and inaccurate witnesses. Investigating officers, in contrast, have the first contact with eyewitnesses, typically shortly after the occurrence of the crime, and they actively interact with them. This may put investigators in a better position than jurors to ascertain the credibility of an eyewitness because they get a first impression before the witness has been contaminated by other factors that might affect his or her behaviour.[1]

Almost all research showing the impact that eyewitnesses have, how their accuracy is overestimated, and how they are believed indiscriminately, has used a mock juror paradigm. An important avenue of research is to examine the impact eyewitnesses have on investigators, since they play a critical role in determining whether an eyewitness's evidence ever makes it to court, and their evaluations of witnesses may influence other aspects of their investigations as well. Yet we know of only two published studies that have focused on the effects of eyewitnesses on investigators (Dahl, Lindsay, & Brimacombe, 2006; Lindsay, Nilsen, & Read, 2000). Dahl et al. (2006) examined how mock investigators were affected by eyewitness identification evidence given by a confederate. Dahl et al. had investigators first interview a confederate-witness about a videotape of a crime that the confederate had allegedly seen. The description provided by the

confederate-witness was scripted to be similar to reports provided by real witnesses in a parallel baseline study using the same video. The investigator was led to believe that the confederate was another student who had also signed up for the study. Based on the description given by the confederate, the investigator chose a suspect from a database of 13 potential suspects; the data base provided a physical description of each potential suspect, along with information pertaining to prior arrest record, current residence and occupation, alibi, etc. The investigator rated the probability that his/her suspect was guilty, and was then shown a photo purportedly of that suspect. The investigator then presented a lineup containing that photo and five foils to the confederate who identified the suspect, identified a foil member, or made no identification. Dahl et al. found that investigators were greatly affected by the identification decision of the confederate-witness even though real witnesses in the baseline study performed at chance on the lineup. Investigators' belief in the guilt of their suspect significantly increased when the confederate identified their suspect and significantly decreased when the confederate identified a foil or rejected the lineup completely.

Dahl et al.'s (2006) use of a confederate witness limits the generalizability and informativeness of their findings (see Clark, Abbe, & Larson, 2006). Their confederate-witness followed a script. Although this script was based on descriptions provided by real witnesses who viewed the same video, it is possible that the fact that the confederate-witnesses had practiced their responses many times created effects similar to trial situations in which witnesses have rehearsed their responses with their lawyers, i.e., they may have inadvertently shown increased confidence (e.g., Boccaccini, Gordon, & Brodsky, 2003). On the other hand, it could be that student-investigators were less over-influenced than they would have been by real witnesses, as the confederate-witness might not have appeared genuine or believable because she was acting a

part. The obvious next step is to conduct a study of mock investigators using real witnesses. Study 1 used real witnesses to examine how student-investigators are affected by eyewitness decisions. Do student-investigators adjust their beliefs accordingly when faced with new information regarding the guilt of their suspect?

Moreover, because Dahl et al. (2006) used a confederate-witness, their findings don't specifically address the issue of whether investigators are equally influenced by accurate and inaccurate witnesses. A second objective of Study 1 was to determine whether student-investigators discriminate between accurate and inaccurate eyewitness identification decisions. As mock jurors don't distinguish between accurate and inaccurate eyewitnesses (Lindsay et al., 1989; Lindsay et al., 1981; Wells et al., 1980), it was predicted that student-investigators would also have difficulty making this distinction.

To test this hypothesis, witnessing conditions were manipulated, so that some witnesses had a better view of the criminal than others. Witnesses who had a good view should provide better descriptions of the criminal than those with a poor view, and should also be in a better position to make an accurate identification. Thus, a third objective of Study 1 was to examine whether investigators put more stock in identification decisions made by witnesses with a good view. Lindsay et al. (1981) manipulated viewing conditions to yield low, moderate, and high levels of identification accuracy and found that although participants were unable to differentiate between accurate and inaccurate witnesses overall, they did adjust their belief according to the witnesses' viewing conditions (although not to the degree they should have). Similarly, Lindsay et al. (2000) found that student-investigators showed more confidence in the accuracy of identification decisions made by witnesses who had a good view compared to witnesses who had a poor view as a result of the increased confidence they showed when making their identification

decisions. Based on these findings, it was anticipated that investigators would be more influenced by witnesses who had a good view than witnesses who had a poor view.

Study 1

Method

Participants

One hundred and forty one pairs of participants were recruited for this study in exchange for optional bonus points in a 100- or 200-level psychology course. Within each pair, one person was randomly assigned to be a witness to a videotaped crime and the other to act as an "investigator" whose job it was to interview the witness about the crime he or she saw and to choose a suspect from a computer database of potential suspects. Each witness was randomly assigned to have either a good or poor view of the criminal and to view either a target present (TP) or a target absent (TA) lineup. In addition, the identification decision that the witness made was included as a grouping variable, separating witnesses who identified the suspect, a foil, or made no identification. Demographic information was not collected from participants, but as random assignment was used any cross-race or gender effects should be scattered across conditions. In the pool from which participants were drawn, the mean age is 22.8 years (*SD* = 3.01 years) and 67.73% are women.

Materials

Videotape of crime. In the poor viewing condition, the video was taped from approximately 20 feet away from a side perspective and was played without sound. It depicted a male culprit committing a robbery of a warehouse along with two male accomplices. The robbery was interrupted by the arrival of a police officer, who arrested the main culprit at the scene of the crime while the two accomplices fled the scene. The video lasts 2 minutes and 5

seconds. The good viewing condition included this same clip but also included an additional minute and ten seconds of exposure to the main culprit at the beginning of the video, including several close-ups of his face.

Police database. The "police database" consisted of a computer program created for this line of research. It included a main page with links to the names of 13 men who could possibly be suspects in the case. When each of these names was clicked, a page displaying information about that man was brought up, including a general physical description (but no picture), birth date, and prior criminal record. Each page also contained a link entitled "additional investigation." If this link was clicked, information about the suspect's vehicle, employment, and alibi for the time of the crime was displayed. In addition, fingerprint information was included for some suspects, given as a percent probability that a fingerprint lifted from the crime scene belonged to the suspect. The culprit resembled one potential suspect ("John Gibbs") better than any other, with the aim of encouraging investigators to select that suspect.

The program required that participants view every suspect's information at least once. Participants could then click on any of the suspects' names to display their information again or were given the option to choose a suspect at this time. When the choose-suspect link was clicked, the names of all 13 suspects were displayed as well as an option to choose none of the suspects. Once participants reached this stage, all decisions were final, i.e., participants could no longer go back to review each suspect's information.

When a suspect was chosen by clicking on his name, a picture purportedly of the suspect was shown. In reality, the computer program showed the same picture regardless of which suspect was chosen. For TP lineups, the suspect was the culprit and for TA lineups the suspect was a similar-looking innocent foil.

When a link on the screen reading “Go to Lineup” was clicked, the photo lineup was displayed. The word “suspect” was displayed below the suspect’s photo, and the words “in jail” appeared below each of the foil’s photos. This was to make it blatantly clear to participants that the other members in the lineup could not possibly have committed the crime. A second link was clicked to show the lineup without these subtitles.

Photo lineup. The photo lineups consisted of 6 frontal head-and-shoulders color photographs arranged in two rows of three. The lineups were constructed using the principles of fair lineup construction so that each person in the lineup matched the same general physical description as the culprit. As well, all members in the lineup were wearing white lab coats in order to eliminate any potential clothing bias effects. The suspect was always in position three in the lineup. TP lineups contained the main culprit and TA lineups contained a similar looking foil.

Questionnaires. Each investigator filled out one questionnaire just prior to conducting the lineup procedure and another thereafter. Among other questions, these questionnaires asked the investigator to indicate the likelihood that the suspect he or she had chosen was guilty, whether he or she would arrest the suspect at this point in the investigation, and, if not, what additional evidence would be required. In addition, witnesses filled out a questionnaire adapted from Bradfield and Wells (1998) that asked them about their viewing conditions and other factors affecting their ability to make an accurate identification. Also, witnesses were asked to indicate how confident they were that they had made a correct identification decision.

Procedure

Each pair of participants arrived and met with the experimenter, who randomly assigned one participant to play the role of the investigator and the other to be the witness. The witness

was then taken into a nearby room where he or she consented to participate in the study and viewed the video depicting the crime. These participants were initially told that they would be asked questions about their perceptions of the video and it was only after watching the video that participant-witnesses were informed of the true nature of the study, i.e., police investigations.

The student-investigator was taken to a second room, where he or she also read and signed a consent form. It was explained to each investigator that the purpose of the study was to simulate aspects of police investigations. Student-investigators were then given instructions for conducting the interview, including an outline of the types of information they should attempt to obtain from the witness (although they were encouraged to ask additional questions as they saw fit).

Once it was clear that the investigator understood the task, the witness was brought back to the room to be interviewed by the investigator about the crime. During the interview, the experimenter took notes on everything the witness said so that the investigator could focus on asking questions and the witness's responses[2]. The investigator was given as much time as needed to conduct the interview.

Once the investigator indicated that he or she was finished with the interview, the witness left the room and the experimenter gave the investigator the notes she had taken along with printed instructions for how to select a suspect from the police database. The purpose of these instructions was not only to explain to investigators how to use the database but also to create a context for them. Investigators were instructed to imagine that they worked for a police department in a small town and that they would be searching a database containing information about potential suspects in the case. They were informed that they would have to go through the complete list of names once, but that they would be able to go back and review people's

information if they needed to before making a decision. As the investigator read over the instructions, the witness was taken back to the other room and told that it would take approximately 10 minutes for the investigator to construct the lineup, and given the option to watch cartoons or play computer games while waiting

The experimenter then returned to the investigator and reviewed the instructions, making sure they were understood by the investigator. The investigator was given sheets to take notes if he or she wished and the computer screen was turned on so that the investigator could begin going through the police database. The investigator went through the list of potential suspects, in most cases ultimately selecting a suspect. Only 3 (2%) investigators failed to choose a suspect, claiming that all of the potential suspects in the police database were poor matches. For those 3 investigators, the study was stopped and both the investigator and the witness were debriefed. For all other investigators, after they'd selected a suspect they filled out the pre-ID questionnaire. Investigators then received an overview of the lineup procedure, which included unbiased instructions, and were instructed not to influence the witness in any way.

After it was clear that the investigator understood the lineup-identification task, the witness was brought back into the room. The investigator conducted the lineup procedure and wrote down the identification decision that the witness made. The investigator was also given an opportunity to ask any additional questions of the witness at this time. The witness was then taken back to the other room and completed a questionnaire, while the investigator completed the post-ID questionnaire. Once the investigator and the witness were finished completing their questionnaires, they were both debriefed and thanked for their participation.

Results

Pre-Identification Questionnaire

Investigators filled out a questionnaire prior to conducting the lineup that asked them to rate the likelihood that their suspect was guilty on a scale from 0% to 100%, whether they would arrest the suspect at this point in the investigation, and what factors had contributed to their selection of their suspect.

Investigator's pre-identification probabilities suspect committed crime. On average, investigators in the TP condition indicated that there was a 60.76% (*SD* = 18.00) chance that their suspect was guilty, and those in the TA condition reported a 61.67% (*SD* = 18.40) chance that their suspect was guilty, $t(136) = .30$, $p = .768$. It is possible that witnesses' behaviour during the initial interview with the investigator varied as a function of the witnesses' subsequent ID decision (e.g., perhaps those who responded correctly on the lineup provided more detailed or confident responses during the initial interview). To address this question, investigators' pre-ID probabilities of guilt were analyzed as a function of the witnesses' subsequent ID-decision. However, pre-identification probabilities did not differ between investigators whose witnesses subsequently made a suspect identification (*M* = 56.43, *SD* = 18.33), a foil identification (*M* = 64.07, *SD* = 19.65), or rejected the lineup (*M* = 61.05, *SD* = 17.04), $F(2, 138) = 1.34$, *MSE* = 438.99, $p = .265$, partial $\eta^2 = .02$.

Decision to arrest. In total, 27.94% (*n* = 19 of 68) of investigators indicated that they would charge the suspect at this point in the TP condition, compared to 34.72% (*n* = 25 of 72) in the TA condition, $z = .68$, $p = .496$. Investigators' mean confidence in their decision to charge or not charge the suspect on a scale from 1 to 10 was 6.05 (*SD* = 1.79) overall. A 2 (Target: Present, Absent) x 2 (Charge Suspect: Yes, No) Analysis of Variance (ANOVA) was conducted to determine whether target presence or the investigators' arrest decision influenced investigators' confidence in that decision. Although the main effect of target presence was not

significant ($F(1, 136) = 1.14$, $MSE = 3.03$, $p = .287$, partial $\eta^2 = .01$, nor was the interaction ($F < 1$), there was a main effect of arrest decision. Investigators who were willing to charge the suspect were significantly more confident in their arrest decision ($M = 6.69$, $SD = 1.42$) than those who were not willing to charge the suspect ($M = 5.75$, $SD = 1.88$), ($F(1, 136) = 8.50$, $MSE = 3.03$, $p = .004$, partial $\eta^2 = .06$. The causal implications of this finding are ambiguous, given that we did not manipulate arrest decision

Ratings of the evidence. Investigators were asked to rate how important physical description, prior record, fingerprint evidence, alibi, and other evidence were in their choice of suspect using percentages so that they summed to 100% across all five types of evidence. A 2 (Target: Present, Absent) x 5 (Evidence: Physical Description, Prior Record, Fingerprint Evidence, Alibi, Other) repeated measures mixed model ANOVA using the Greenhouse-Geisser correction to account for violations of the assumption of sphericity indicated that while target presence did not affect investigators' ratings of the evidence and the interaction was not significant (Fs < 1), there were significant differences in rated importance between the evidence factors, $F(3.03, 417.52) = 103.32$, $MSE = 313.66$, $p < .001$, partial $\eta^2 = .43$. Subsequent pair-wise comparisons were carried out using the Bonferroni correction. Physical description ($M = 42.53$, $SD = 20.40$) was rated as significantly more important than prior criminal record ($M = 17.11$, $SD = 16.28$), $t(139) = 12.09$, $p < .001$, Cohen's $d = 1.38$, fingerprints ($M = 22.22$, $SD = 16.65$), $t(139) = 8.41$, $p < .001$, Cohen's $d = 1.09$, alibi ($M = 20.20$, $SD = 15.34$), $t(139) = 10.33$, $p < .001$, Cohen's $d = 1.24$, and other evidence ($M = 6.11$, $SD = 10.86$), $t(139) = 17.46$, $p < .001$, Cohen's $d = 2.23$. Prior criminal record was rated as significantly more important than other evidence, $t(139) = 6.86$, $p < .001$, Cohen's $d = 0.80$. Fingerprints were rated as significantly more important than prior criminal record and other evidence ($t(139) = 3.49$, $p = .001$, Cohen's d

= 0.31 and $t(139) = 9.37$, $p < .001$, Cohen's $d = 1.15$ respectively). Finally, alibi was rated as significantly more important than other evidence, $t(139) = 8.70$, $p < .001$, Cohen's $d = 1.06$.

Witnesses' Identification Decisions

Identification decisions for all 141 witnesses by viewing condition and target presence can be seen in Table 1. Although there were subtle differences in accuracy depending on whether the witness had a good or a poor view[3], the viewing quality manipulation did not significantly affect accuracy overall ($X^2(1) = 2.38$, $p = .123$) and so good and poor viewing conditions were collapsed for all analyses. Comparing suspect choices alone, it can be seen that witnesses performed poorly on the lineup, as the suspect was almost as likely to be chosen from a TA lineup (suspect was chosen 13.70% of the time) as a TP lineup (suspect was chosen 19.12% of the time), $z = .86$, $p = .390$. A similar pattern occurred for lineup rejections, with 53.42% correctly rejecting the TA lineup compared to 51.47% incorrectly rejecting the TP lineup, $z = .22$, $p = .823$.

Post-Identification Questionnaire

After administering the lineup, investigators filled out a second questionnaire on which they were again asked to rate the likelihood that their suspect was guilty on a scale from 0% to 100% and whether they would arrest the suspect at this point in the investigation. Information gain analyses are also reported indicating how much investigators shifted in their belief that their suspect was guilty compared to how much they should have based on the accuracy of witnesses on the lineup. In addition, the question of whether investigators were able to discriminate between accurate and inaccurate witnesses is addressed.

Post-identification probability suspect committed crime. A 2 (Phase: Pre-ID, Post-ID) x 3 (ID Decision: ID Suspect, ID Foil, Not Present) repeated measures mixed model ANOVA[4] was

used to investigate whether there were significant differences in participants' estimated probability that the suspect was the criminal. The interaction was significant, $F(2, 135) = 69.46$, $MSE = 234.23$, $p < .001$, partial $\eta^2 = .51$. Paired t-tests revealed that when the suspect was identified, investigators' guilt probabilities went up significantly from pre-lineup ($M = 56.43$, $SD = 18.33$) to post-lineup ($M = 83.04$, $SD = 14.48$), $t(22) = -8.09$, $p < .001$, Cohen's $d = -1.61$. If the lineup was rejected, investigators' guilt probabilities went down significantly from pre-lineup ($M = 61.05$, $SD = 17.04$) to post-lineup ($M = 27.43$, $SD = 19.78$), $t(71) = 13.03$, $p < .001$, Cohen's $d = 1.82$. Finally, when a foil was identified, guilt probabilities also dropped significantly from pre-lineup ($M = 64.07$, $SD = 19.65$) to post-lineup ($M = 37.26$, $SD = 21.14$), $t(42) = 7.41$, $p < .001$, Cohen's $d = 1.31$.

There was a main effect of phase, $F(1, 135) = 30.35$, $MSE = 234.23$, $p < .001$, partial $\eta^2 = .18$. This effect reflected the fact that most witnesses (84%) made an exculpatory identification judgment. There was also a main effect of identification decision, $F(2, 135) = 23.90$, $MSE = 475.00$, $p < .001$, partial $\eta^2 = .26$. However, because the identification decision only affected post-identification probabilities, post hoc tests were only conducted on the post-identification ratings.

Post hoc analyses using the Bonferroni correction revealed that investigators were significantly more likely to think their suspect was the criminal when the witness identified their suspect ($M = 83.04$, $SD = 14.48$) than when the witness identified a known-innocent foil ($M = 37.26$, $SD = 21.14$) ($t(65) = 10.43$, $p < .001$, Cohen's $d = 2.53$) or rejected the lineup ($M = 27.43$, $SD = 19.78$) ($t(95) = 14.65$, $p < .001$, Cohen's $d = 3.21$). In addition, investigators were significantly more likely to think their suspect was the culprit if an identification of a foil was made than if the lineup was rejected, $t(116) = 2.50$, $p = .014$, Cohen's $d = 0.48$. These results

can be seen in Figure 1.

Arrest decision. In the ID-suspect condition, 60.00% (*n* = 6 of 10) reported they would arrest the suspect in the TA condition compared to 76.92% (*n* = 10 of 13) in the TP condition, *z* = .42, *p* = .677. In the ID-foil condition, 30.00% (*n* = 6 of 20) indicated they would arrest the suspect in the TA condition compared to 34.78% (*n* = 8 of 23) in the TP condition, *z* = .84, *p* = .400. Finally, in the no-ID condition 7.69% (*n* = 3 of 39) indicated they would arrest the suspect in the TA condition whereas 14.29% (*n* = 5 of 35) indicated they would arrest the suspect in the TP condition, z = .53, p = .591. Relative to the pre-ID decisions to arrest, the frequency of investigators choosing to arrest decreased significantly in the no-ID condition (*z* = 2.75, *p* = .006) and increased significantly in the ID-suspect condition (*z* = 3.21, *p* = .001) but did not change significantly from pre to post-ID in the ID-foil condition (*z* = .65, *p* = .513).

Information gain. How much should investigators have been influenced by each type of identification decision in this study? To determine this, information gain analyses were conducted using the equations from Wells and Olson (2002) which are included in the Appendix. These equations, which are based on a Bayesian analysis, hold that the amount of information gained by an identification decision is equal to the absolute value of the prior probability that the suspect is the culprit minus the posterior probability that the suspect is the culprit given the identification decision of the witness. Each investigator's pre-identification probability that the suspect was the culprit was used as the prior probability that the culprit was in the lineup.

Circles are included on Figure 1 indicating how much investigators should have shifted their beliefs based on the identification decision of the witness. It can be seen that investigators were unduly influenced by all identification decisions. When the suspect was identified, investigators returned significantly higher probabilities of his guilt (*M* = 83.04, *SD* = 14.48) than

they should have based on the information gained from the identification decision (M = 63.42, SD = 17.11), $t(22) = 6.22$, $p < .001$. Conversely, when a foil was identified or the lineup was rejected, investigators returned significantly lower probabilities of guilt (Foil ID: M = 37.26, SD = 21.14, No ID: M = 27.40, SD = 19.91) than they should have based on the information gained from these identification decisions (Foil ID: M = 61.90, SD = 19.83, No ID: M = 60.27, SD = 17.11), $t(42) = -6.86$, $p < .001$ and $t(71) = -12.70$, $p < .001$ respectively.

Did investigators discriminate between accurate and inaccurate witnesses?

Investigators' percent probabilities that the suspect committed the crime were compared for those whose witness had made a correct identification decision versus an incorrect identification decision. As can be seen in Figure 2, investigators' post-identification ratings of the guilt of the suspect were equal for suspect identifications regardless of whether the identification was correct (M = 83.85, SD = 14.95) or incorrect (M = 82.00, SD = 14.57), $t(21) = .30$, $p = .770$, Cohen's d = 0.13. Likewise, investigators' post-identification ratings of the suspect's guilt did not significantly differ when the lineup was rejected whether the decision was correct (M = 25.59, SD = 17.99) or incorrect (M = 30.01, SD = 22.14), $t(71) = -.95$, $p = .345$, Cohen's d = -0.22. This is not surprising given how close to chance witnesses were; most correct responses were likely to be guesses.

Eyewitness Questionnaires

Eyewitnesses also filled out a questionnaire on which they made a number of ratings about their basis for making an identification. Specifically, they were asked to rate their quality of view, length of exposure to the criminal and ability to make out his face, distance from the criminal, level of attention paid to the video, confidence in their ID decision, difficulty of and length of time to make their ID decision, willingness to testify in court, and whether another

eyewitness with the same view should be trusted. A multivariate analysis of variance (MANOVA) [5] was used to examine whether accurate witnesses differed from inaccurate witnesses on any of the ratings. The MANOVA was not significant, Pillai's Trace = .12, $F(11, 122) = 1.53$, $p = .130$.

Was eyewitness confidence correlated with other variables?

Both investigators and the witnesses themselves rated the witnesses' confidence in their identification decision on scales from 1 to 10 and 1 to 7 respectively. To examine the relationship between the witnesses' confidence and their impact on investigators, impact scores were calculated as the difference between investigators' pre-identification ratings of the probability that their suspect was guilty versus their post-identification ratings. Correlations were then calculated to see whether there was a relationship between eyewitnesses' self-reported confidence and these impact scores, calculated separately for each identification decision. Eyewitness confidence was not significantly correlated with the impact that eyewitnesses had on investigators when a foil was identified or the lineup rejected (No ID: Pearson's $r = -.10$, $p = .432$; ID Foil: Pearson's $r = -.05$, $p = .754$). However, the correlation between eyewitness confidence and eyewitness impact was marginally significant when the suspect was identified (Suspect ID: Pearson's $r = .39$, $p = .069$). The more confident the witness was, the greater the increase in investigators' belief in the guilt of the suspect after he was identified.

The impact scores were then correlated with the perceived confidence of the witness for each type of identification made to see whether witnesses' confidence as perceived by investigators was related to the impact that eyewitnesses had. What may matter is not how confident the witness is but how confident the investigator thinks the witness is A marginally significant correlation was found for witnesses who had rejected the lineup, Pearson's $r = -.24$, p

= .052: The more confident a witness was perceived to be, the greater the drop in investigators' ratings of guilt after no identification was made. The other correlations did not approach significance (Suspect ID: Pearson's r = .21, p = .345, Foil ID: Pearson's r = .05, p = .733).

Discussion

Investigators were strongly influenced by witnesses' identification decisions. If the suspect was identified, investigators' ratings of the suspect's guilt rose substantially. Conversely, if a foil was identified or the lineup was rejected, guilt ratings dropped considerably. Overall, investigators were significantly more likely to think their suspect was the criminal if he had been identified than if a foil was identified or the lineup was rejected. Investigators were also significantly more likely to think their suspect was guilty if a foil ID was made than if the lineup was rejected. These results replicate findings by Dahl et al. (2006) that tested the impact of confederate-witnesses' identification decisions on investigators and indicates that these results generalize to the impact of real witnesses on investigators.

Information gain analyses indicated that people were more swayed by eyewitness identification decisions than they should have been based on how informative each type of identification decision (suspect ID, foil ID, no ID) was as to the guilt of the suspect. With respect to witnesses' performance on the lineup, it was found that witnesses were almost equally likely to reject a lineup containing the culprit as one that did not contain the culprit. Likewise, witnesses were almost as likely to choose innocent suspects as guilty suspects. Witnesses' poor performance precluded meaningful tests of investigators' discrimination between witnesses with a good or a poor view or between accurate and inaccurate witnesses.

Rationale for Study 2

In Study 1 it was expected that those who had a good view of the criminal would be more accurate than those who had a poor view of the criminal. However, performance did not differ significantly between the two groups of witnesses, as they both performed poorly, with correct identifications roughly at chance levels. Although witnesses in the good view condition received an additional 70 seconds of exposure to the culprit, the majority of this exposure was from at least 20 feet away. As well, our subjective impression is that the culprit looks different in the lineup photo than in the video. Consistent with that impression, informal pilot testing suggested that people perform poorly on this lineup even if the video is playing while they are making their lineup decisions. As the manipulation was too weak to lead to differences in accuracy rates, it was not possible to examine the effects of viewing conditions in Study 1. In Study 2, new materials were used in which the additional exposure time to the culprit was shot head-on from only a few feet away. These videos were pilot tested to ensure that they did lead to high and low accuracy rates.

Another issue in Study 1 relates to the fact that even had the good and poor viewing conditions led to high and low accuracy rates, very few selections of the innocent suspect occurred, which made it necessary to collapse across good and poor viewing conditions to be able to conduct meaningful analyses. In the current study, pilot testing was used to determine the most similar-looking person to the culprit in our lineup. This person was then used as the innocent suspect in order to increase the frequency with which identifications of the innocent suspect occurred.

Given that accuracy in Study 1 was so low, in Study 2, an incentive was introduced to increase participants' motivation to be accurate. As an added benefit, the incentive may increase the generalizeability of Study 2 to the real world where both witnesses and investigators are

highly motivated to catch the guilty person. The incentive consisted of a chance to be part of two separate $100 draws for witnesses and investigators who made accurate decisions in terms of identification and arrest decisions.

Method

Participants

One hundred and eighty pairs of University students were recruited for this study and received either course credit or five dollars for their participation. The general procedure was the same as Study 1 in that within each pair, one person was randomly assigned to be a witness to a videotaped crime and the other to act as an "investigator." Likewise each witness was randomly assigned to have either a good or poor view of the criminal and to view either a TP or a TA lineup. The identification decision that the witness made was once again included as a grouping variable separating participants who identified the suspect, a foil, or made no identification. Participation took approximately 45 minutes.

Materials

Videotape of crime. In the poor viewing condition, the video was taped at a distance of approximately 10 feet and was set in the common area of a building on campus. The video depicted a female culprit stealing two wallets from a male's bag and a female's jacket after they left their things unattended. The video lasted 2 minutes and 3 seconds. The good viewing condition included this same clip but also included an additional minute and thirty seconds of exposure to the culprit at the beginning of the video, including several close-ups of her face.

Police database. The "police database" consisted of a computer program analogous to that used in Study 1 but adapted for the current study by including information about 13 women who could be possible suspects in the case. As in Study 1, the information in the database fit one

potential suspect by the name of Jane Gibbs, whose description was created based on pilot testing.

Photo lineup. The photo lineup consisted of 6 frontal head-and-shoulders photographs arranged in two rows. The lineups were constructed using the principles of fair lineup construction so that each person in the lineup matched the same general physical description as the culprit. As in Study 1, all members in the lineup wore white lab coats to eliminate any potential clothing bias effects and the suspect was always in position three in the lineup. Present lineups contained the culprit and absent lineups contained a similar looking foil, chosen based on pilot testing.

Questionnaires. Investigators and witnesses filled out the same questionnaires as in Study 1.

Procedure

The procedure was identical to that in Study 1 except for one difference. Investigators and witnesses in the current study were informed that those who made correct decisions (i.e., investigators who correctly chose to arrest or not arrest their suspect and witnesses who made an accurate identification decision) would be placed in a draw for $100. In the current study, only one investigator failed to choose a suspect indicating that no suspect was a good match; as in Study 1, the study was stopped and both the investigator and the witness were debriefed.

Results

Pre-Identification Questionnaire

As in Study 1, investigators filled out a questionnaire prior to conducting the lineup that asked them to rate the likelihood that their suspect was guilty on a scale from 0% to 100%,

whether they would arrest the suspect at this point in the investigation, and what factors had contributed to their selection of their suspect.

Investigator's pre-identification probabilities suspect committed crime. Overall, investigators indicated that there was a 64.39% (SD = 17.87) chance that their suspect was guilty. A 2 (View: Good, Poor) x 2 (Target: Present, Absent) ANOVA found that there were no differences in investigators' pre-identification probabilities that their suspect was guilty as a function of the witness's view or target presence (*Fs* < 1). In addition, pre-identification probabilities did not differ between investigators whose witnesses subsequently made a suspect identification (*M* = 65.48, *SD* = 17.58), a foil identification (*M* = 64.61, *SD* = 19.44), or rejected the lineup (*M* = 63.68, *SD* = 18.14), *F* < 1. These results replicate Study 1.

Decision to arrest. In the poor view condition, 32.94% (TP: 9 of 23, TA: 8 of 26, *z* = .754) of investigators indicated that they would charge the suspect at this point, compared to 28.42% (TP: 9 of 34, TA: 15 of 43, *z* = .54, *p* = .587) of investigators in the good condition, *z* = .65, *p* = .516. Investigators' were asked to rate their confidence in their arrest decision on a scale from 1 to 10. Investigators' mean confidence in their decision to charge or not charge the suspect on a scale from 1 to 10 was 6.45 (*SD* = 1.64). A 2 (View: Good, Poor) x 2 (Target: Present, Absent) x 2 (Charge Suspect: Yes, No) ANOVA was conducted to determine whether target presence or arrest decision influenced investigators' confidence in that decision. None of the main effects or interactions was significant so these results will not be discussed further (all *F*s < 1.2 except Charge Suspect: $F(1, 172) = 3.26$, $MSE = 2.69$, $p = .073$, partial $\eta^2 = .02$). These results replicated Study 1, other than the finding that in Study 1, rather than approaching significance, investigators were significantly more confident in their arrest decision if they chose to charge the suspect.

Ratings of the evidence. Participants were asked to rate how important physical description, prior record, fingerprint evidence, alibi, and other evidence were in their choice of suspect using percentages so that they summed to 100% across all five types of evidence. A 2 (View: Good, Poor) x 2 (Target: Present, Absent) x 5 (Evidence: physical description, prior record, fingerprint evidence, alibi, other evidence) repeated measures mixed model ANOVA was used to investigate whether viewing conditions or target presence had any effect on investigators' rated importance between the evidence factors. Using the Greenhouse-Geisser correction to account for violations of the assumption of sphericity, it was found that there were significant differences in rated importance between the evidence factors overall, $F(2.42, 409.66) = 173.46$, $MSE = 365.64$, $p < .001$, partial $\eta^2 = .51$. Subsequent pair-wise comparisons were carried out using the Bonferroni correction. Physical description ($M = 42.62$, $SD = 19.15$) was rated as significantly more important than fingerprints ($M = 25.37$, $SD = 17.05$), $t(177) = 7.33$, $p < .001$, Cohen's $d = 0.95$, alibi ($M = 16.33$, $SD = 10.37$), $t(176) = 14.31$, $p < .001$, Cohen's $d = 1.71$, prior criminal record ($M = 15.89$, $SD = 12.50$), $t(176) = 14.77$, $p < .001$, Cohen's $d = 1.65$, and other evidence ($M = 2.59$, $SD = 6.46$), $t(172) = 25.92$, $p < .001$, Cohen's $d = 2.80$. Fingerprints were rated as significantly more important than alibi, prior criminal record, and other evidence ($t(176) = 6.25$, $p < .001$, Cohen's $d = 0.64$, $t(176) = 5.49$, $p < .001$, Cohen's $d = 0.63$, and $t(172) = 15.99$, $p < .001$, Cohen's $d = 1.77$ respectively). Finally, alibi and prior criminal record were rated as significantly more important than other evidence, $t(173) = 15.01$, $p < .001$, Cohen's $d = 1.59$ and $t(173) = 13.03$, $p < .001$, Cohen's $d = 1.34$. There was no main effect of viewing conditions or target presence, nor were any of the interactions significant (all Fs < 1, except for the View x Target x Evidence interaction, $F(2.42, 409.66) = 1.47$, $MSE = 365.64$, $p = .227$, partial $\eta^2 = .01$). Once again, these results replicate Study 1.

Witnesses' Identification Decisions

Identification decisions for all 180 witnesses by viewing condition and target presence can be seen in Table 2. Witnesses who had a good view of the criminal were significantly more accurate than witnesses who had a poor view of the criminal, making accurate decisions 62.11% of the time (n = 59 of 95) and 30.59% of the time (n = 26 of 85) respectively, $z = 4.44$, $p < .001$. In the poor view condition, witnesses were almost as likely to chose the suspect from an absent lineup (17.78% of the time, n = 8 of 45) as a present lineup (22.50% of the time, n = 9 of 40), $z = .53$, $p = .595$, whereas in the good view condition, witnesses were significantly more likely to choose the suspect in the present lineup (60.98% of the time, n = 25 of 41) than the absent lineup (14.82% of the time, n = 8 of 54), $z = 5.06$, $p < .001$. A similar pattern occurred for lineup rejections whereby in the poor view condition witnesses were only slightly more likely to reject an absent lineup (40.00% of the time, n = 18 of 45) than a present lineup (35.00% of the time, n = 14 of 40), $z = .47$, $p = .638$, whereas in the good view condition, witnesses were significantly more likely to reject an absent lineup (64.81% of the time, n = 35 of 54) than a present lineup (21.95% of the time, n = 9 of 41), $z = 4.62$, $p < .001$.

Errors in the Good View condition were fairly rare, which led us to test more participants in the Good View condition (95) than in the Poor View condition (85). Even so, only 7 witnesses in the Good View condition tested on the TP lineup identified a foil. Perhaps partly because of this small cell size, data for witnesses who made foil IDs differed from those of prior studies and were difficult to interpret. In real-world cases, foil IDs are of limited interest because they are known errors. We therefore do not report these data here, but they are available from the first author.

Post-Identification Questionnaire

After administering the lineup, investigators filled out a second questionnaire on which they were again asked to rate the likelihood that their suspect was guilty on a scale from 0% to 100% and whether they would arrest the suspect at this point in the investigation. As in Study 1, information gain analyses are also reported indicating how much investigators shifted in their belief that their suspect was guilty compared to how much they should have based on the accuracy of witnesses on the lineup. In addition, the question of whether investigators were able to discriminate between accurate and inaccurate witnesses is addressed.

Post-identification probability suspect committed crime. A 2 (Phase: Pre-ID, Post-ID) x 2 (View: Good, Poor) x 2 (Decision: ID Suspect, Not Present) repeated measures mixed model ANOVA was used to investigate whether there were significant differences in investigators' estimated probability that the suspect was the criminal as a function of the quality of the witnesses' view.[6] The phase x ID decision interaction was significant, $F(1, 122) = 215.67$, $MSE = 192.85$, $p < .001$, partial $\eta^2 = .64$). Paired t-tests revealed that when the suspect was identified, investigators' guilt probabilities rose significantly from pre-lineup ($M = 65.48$, $SD = 17.58$) to post-lineup ($M = 85.60$, $SD = 11.62$), $t(49) = -9.56$, $p < .001$, Cohen's $d = -1.35$. If the lineup was rejected, investigators' guilt probabilities dropped significantly from pre-lineup ($M = 63.68$, $SD = 18.14$) to post-lineup ($M = 29.30$, $SD = 22.92$), $t(75) = 13.53$, $p < .001$, Cohen's $d = 1.66$.

There was a main effect of phase, $F(1, 122) = 12.96$, $MSE = 192.85$, $p < .001$, partial $\eta^2 = .10$. This effect reflected the fact that, as in Study 1, the majority of witnesses (60%) made an exculpatory identification judgment. There was also a main effect of identification decision, $F(1, 122) = 90.58$, $MSE = 507.58$, $p < .001$, partial $\eta^2 = .43$. Identification decision affected post-identification probabilities ($t(124) = -16.05$, $p < .001$, Cohen's $d = 3.10$) but not pre-identification probabilities ($t(124) = -.55$, $p = .582$, Cohen's $d = -0.10$). Investigators were

significantly more likely to think their suspect was the criminal when their witness identified their suspect (M = 85.60, SD = 11.62) than when the witness rejected the lineup (M = 29.30, SD = 22.92). The main effect of view was not significant nor was the view x ID decision interaction (both Fs < 1). The overall results can be seen in Figure 3. These results are similar to Study 1.

Decision to arrest. Investigators were again asked whether they would arrest the suspect based on all the information that they had obtained thus far. In the ID-suspect condition, 97.00% (TP: 24 of 25, TA: 8 of 8, z = .61, p = .542) reported they would arrest the suspect when witnesses had a good view, which was not significantly different than the 88.24% (TP: 8 of 9, TA: 7 of 8, z = .67, p = .506) who reported that they would arrest the suspect when witnesses had a poor view, z = .57, p = .307. Likewise in the no-ID condition, the 13.6% (TP: 0 of 9, TA: 6 of 35, z = 1.12, p = .428) who indicated that they would arrest the suspect when witnesses had a good view was not significantly different than the 9.38% (TP: 0 of 14, TA: 3 of 18, z = .99, p = .320) who indicated that they would arrest the suspect when the witness had a poor view, z = .08, p = .571. Compared to the pre-ID decisions to arrest, the frequency of investigators choosing to arrest increased significantly from 36.00% (n = 18 of 50) to 94.00% (n = 47 of 50) in the ID-suspect condition (z = 7.58, p < .001) and decreased significantly from 30.26% (n = 23 of 76) to 11.84% (n = 9 of 76) in the no-ID condition (z = 2.85, p = .004). This finding replicates Study 1.

Information gain. To determine whether investigators adjusted their beliefs appropriately based on each type of identification decision for this particular lineup, information gain analyses were conducted as outlined in Study 1 using the equations from Wells and Olson (2002) (see Appendix). Circles on Figure 3 indicate how much investigators should have shifted their estimates based on the identification decision of the witness. Paired t-tests revealed that in the poor view condition, the results replicated Study 1 in that investigators were unduly influenced

by all identification decisions, returning higher probabilities of guilt when the suspect was identified (M = 85.06, SD = 14.39) and lower probabilities of guilt when the lineup was rejected (M = 30.91, SD = 24.27) than they should have based on the information gained by these identification decisions (ID-Suspect: M = 67.79, SD = 15.09; No-ID: M = 59.50, SD = 18.29), $t(16) = 5.65$, $p < .001$ and $t(31) = -8.31$, $p < .001$ respectively. These findings are similar to Study 1 in which participants were over-influenced by all identification decisions based on the poor performance of witnesses on the lineup.

However, in the good view condition, a different pattern emerged. When the lineup was rejected, investigators were over-influenced similar to the poor view condition, lowering their probability of guilt ratings (M = 28.14, SD = 22.10) more than they should have based on the accuracy of the witnesses (M = 41.83, SD = 17.49), $t(43) = -3.70$, $p = .001$. When the suspect was identified, in contrast, investigators were highly calibrated, shifting their belief in the guilt of the suspect approximately as much (M = 85.88.14, SD = 10.15) as they should have based on the accuracy of the witnesses (M = 87.07, SD = 12.29), $t(31) = -.59$, $p = .558$. As discussed in greater detail in the General Discussion, we speculate that if conditions led eyewitnesses to be extremely accurate, such that information gain was very high, investigators would be under-influenced by them.

Were investigators able to discriminate between accurate and inaccurate witnesses? Separate t-tests were conducted examining whether investigators were more swayed by suspect identifications and lineup rejections that were correct than those that were incorrect. Investigators' post-identification ratings of the guilt of the suspect were not significantly different when the lineup was rejected regardless of whether the identification was correct or incorrect for either the good (correct: M = 28.23, SD = 23.39, incorrect: M = 27.78, SD = 17.34)

or poor viewing conditions (correct: M = 36.11, SD = 25.06, incorrect: M = 24.11, SD = 22.31, $t(42)$ = -.054, p = .957, Cohen's d = 0.02 and $t(30)$ = -1.40, p = .173, Cohen's d = 0.17 respectively). However, although the accuracy of the identification decision did not affect investigators' post-identification ratings when the suspect was identified in the poor viewing condition (correct: M = 85.00, SD = 19.82, incorrect: M = 85.13, SD = 8.98, $t(15)$ = .02, p = .986, Cohen's d = -0.01), which is similar to Study 1, in the good viewing condition it was found that investigators returned significantly higher post-identification probabilities when the witness was accurate than when the witness was not (correct: M = 88.56, SD = 9.02, incorrect: M = 77.50, SD = 9.26, $t(31)$ = -3.00, p = .005, Cohen's d = 1.21). These results can be seen in Figure 4.

Eyewitness Questionnaires

Eyewitnesses also filled out a questionnaire on which they made a number of ratings about their bases for making an identification. As indicated in Study 1, these ratings related to their quality of view, length of exposure to the criminal and ability to make out his face, distance from the criminal, level of attention paid to the video, confidence in their ID decision, difficulty of and length of time to make their ID decision, willingness to testify in court, and whether another eyewitness with the same view should be trusted. A MANOVA was conducted to see whether witness accuracy (correct, incorrect) or viewing conditions (poor view, good view) were related to these ratings. Although the main effect of accuracy and the interaction were not significant (Fs < 1), there was a main effect of view, Pillai's Trace = .32, $F(12, 73)$ = 2.80, p = .003. Compared to witnesses who had a poor view, witnesses who had a good view reported seeing the criminal's face for a significantly greater number of seconds ($F(1, 84)$ = 8.35, MSE = 8761.90, p = .005, partial η^2 = .09), being able to make out the criminal's face better ($F(1, 84)$ = 18.37, MSE = 1.77, p < .001, partial η^2 = .18), being a fewer number of feet away from the

criminal ($F(1, 84) = 10.35$, $MSE = 7.29$, $p = .002$, partial $\eta^2 = .11$), feeling more confident they would be able to identify the criminal ($F(1, 84) = 7.99$, $MSE = 2.10$, $p = .006$, partial $\eta^2 = .09$), providing a higher rating of the probability their identification decision was correct ($F(1, 84) = 7.38$, $MSE = 1138.77$, $p = .008$, partial $\eta^2 = .08$), being more willing to testify in court ($F(1, 84) = 7.85$, $MSE = 5.31$, $p = .006$, partial $\eta^2 = .09$), and having a better basis to make an identification ($F(1, 84) = 7.05$, $MSE = 4.14$, $p = .009$, partial $\eta^2 = .08$).

Was eyewitness confidence correlated with other variables?

Both the investigators and the witnesses themselves rated the witnesses' confidence in their identification decision on scales from 1 to 10 and 1 to 7 respectively. Difference scores were created by subtracting the pre-ID probabilities of guilt from the post-ID probabilities of guilt and were correlated with eyewitnesses' self-reported confidence and investigators' perceived confidence of the witness for each type of identification made. Although there was no relationship between eyewitnesses' self-reported confidence and their impact on investigators (Suspect ID: Pearson's $r = .14$, $p = .346$; No ID: Pearson's $r = -.16$, $p = .166$), it was found that the confidence of the witness as perceived by investigators was associated with a significantly greater drop in investigators' ratings of guilt after no identification was made (Pearson's $r = -.49$, $p < .001$) and a significantly greater increase in investigators' ratings of guilt after the suspect was identified (Pearson's $r = -.32$, $p = .030$). These findings are similar to Study 1.

Discussion

The viewing condition manipulation was successful in the current study as witnesses were more than twice as likely to be accurate when they had a good view of the criminal as when they had a poor view. Witnesses' performance in the poor viewing condition was at chance which is similar to witnesses' performance in Study 1.

Investigators' post-identification probabilities replicated our past research in that when the suspect was identified, post-identification probabilities of guilt rose substantially, whereas if the lineup was rejected, investigators' post-identification probabilities fell. Comparing across the ID decisions, it was also found that post-identification probabilities were significantly higher if the suspect was identified than if no identification was made. However, contrary to our prediction, investigators did not put more stock into identification decisions made by witnesses with a good view.

Information gain analyses replicated Study 1 in the poor view condition, that is, investigators put too much stock in witnesses' identification decisions by adjusting their belief in the guilt of their suspect more than was warranted based on the identification accuracy of the witnesses. However, in the good view condition, although investigators were over-influenced by a lineup rejection (i.e., they decreased their belief in the guilt of the suspect more than they should have based on the accuracy of the witnesses), they were very well calibrated when the suspect was identified. It may be that subject-investigators are not universally over-influenced by witnesses' ID judgments, but rather that they (a) tend to be quite strongly influenced and (b) are insensitive to factors associated with variations in witnesses' accuracy; thus when witnesses' accuracy is low to moderate they tend to be over-influenced, when it is fairly high they are appropriately influenced, and when it is very high they may be under-influenced. Further research is needed to explore these speculations.

Investigators did not discriminate between accurate versus inaccurate witnesses except in one circumstance. When witnesses who had a good view identified the suspect, investigators were more likely to believe accurate witnesses than inaccurate witnesses. Although these findings contrast with past research indicating that mock jurors are unable to discriminate

between accurate and inaccurate witnesses (e.g., Lindsay, Wells, & O'Connell, 1989), they are consistent with findings by Lindsay et al. (2000) that student investigators had more confidence in accurate than inaccurate witnesses.

Although investigators for the most part did not discriminate between accurate and inaccurate witnesses or those who had a good or a poor view, the witnesses themselves did report differences. Witnesses who were exposed to a good view of the criminal reported having a longer view of the criminal's face, being able to make out the criminal's face better, being closer to the criminal, being more confident that they could identify the criminal prior to viewing the lineup, that there was a higher probability that their decision was accurate, being more willing to testify in court, and having a stronger basis to make an identification. Witnesses were clearly sensitive to their viewing conditions and the impact that it would have on their ability to make an accurate identification even if investigators were unable to pick up on this difference.

General Discussion

This research had two central objectives. The first objective was to examine how subject investigators are influenced by real witnesses and the second objective was to assess the extent to which these investigators are sensitive to witnesses' accuracy and to variables that modulate witnesses' accuracy. With respect to the first objective, it was found that subject investigators were greatly influenced by real witnesses' identification decisions. When witnesses were highly accurate (as was the case when they had a good view) then the impact they had on investigators was justified and investigators made decisions accordingly. However, when witnesses' accuracy was compromised (as was the case in Study 1, and when witnesses had a poor view in Study 2), investigators placed undue emphasis on witnesses' identification decisions. Thus, with respect to objective two, it was found that witnesses were for the most part unable to distinguish between

accurate and inaccurate eyewitnesses as they tended to weight decisions from both equally. The one exception was that when witnesses had a good view, accurate witnesses had more impact than inaccurate witnesses on investigators. Finally, investigators showed virtually no sensitivity to witnesses' viewing conditions, in spite of this factor having a dramatic effect on witnesses' accuracy.

Although it was not the main focus of our research, these studies also contribute to a small body of evidence on the effects of exposure duration on subsequent lineup performance. Indeed, to the best of our knowledge these are the only experimental manipulations of exposure duration including a designated suspect in TA lineups[7]. The results of Study 1 indicate that exposure duration can have null effects if the sensory quality of the added information is weak and/or the match between the culprit's appearance during the crime and during the lineup is poor. Study 2, in contrast, showed that increased exposure to close-up views can substantially increase suspect IDs in TP lineups and substantially decrease suspect IDs in TA lineups.

One unexpected finding was common to both studies. In both studies there were no differences in investigators' pre-identification probabilities of guilt depending on whether the witness had been exposed to good or poor viewing conditions. Although this is not surprising in the first study in which performance did not differ between the two groups, we expected that in the second study the information provided by witnesses in the poor viewing condition would be less detailed than that related by witnesses in the good viewing condition, potentially leading investigators to be less certain in their choices. There was no evidence of such an effect. An interesting avenue for future research would be to videotape the interviews and analyze the behaviour of the witnesses to determine whether there are any observable differences in witnesses' behaviour depending on their viewing conditions.

Although the major findings of this research replicated findings by Dahl et al. (2006), there were a couple of differences that should be addressed. In terms of investigators' suspect selections, there was substantially more variability in the current research than in Dahl et al.'s research using confederate witnesses. Whereas all of Dahl et al.'s investigators chose John Gibbs as the witness, only half did in Study 1 (84% chose Jane Gibbs in Study 2). This difference is believed to be because the descriptions given by the witnesses were genuine, rather than scripted as in Dahl et al.'s study. Although confederate-witnesses' responses in Dahl et al.'s research were based on descriptions given by actual witnesses, it appears there are differences between witnesses in terms of their accuracy and detail that Dahl et al.'s confederate-witnesses did not capture. If all witnesses had given detailed and accurate descriptions of the criminal they had seen investigators would have invariably been led to John Gibbs as their suspect. Investigators who chose John Gibbs did so presumably because they interviewed better witnesses. This is further supported by findings that investigators who chose John Gibbs as their suspect provided higher pre-identification guilt ratings than investigators who chose other suspects (John Gibbs: $M = 67.16$, $SD = 14.04$, Others: $M = 55.10$, $SD = 19.88$, $t(120.23) = 4.11$, $p < .001$, Cohen's $d = .70$).

A second difference is that investigators in the current research provided significantly lower pre-identification probabilities that their suspect was guilty ($M = 61.20\%$, $SD = 18.15\%$ for Study 1) than investigators in Dahl et al.'s (2006) research ($M = 77.08\%$, $SD = 9.08$), $t(163) = 6.60$, $p < .001$[8]. This difference is likely also related to individual differences in witnesses' ability to accurately describe the criminal. Even investigators who chose John Gibbs as their suspect and thus likely interviewed better witnesses provided a lower mean pre-identification probability that their suspect was guilty ($M = 67.16\%$, $SD = 14.04\%$) than investigators in Dahl

et al.'s research ($t(62) = -3.97$, $p < .001$) indicating that even good witnesses were less compelling than Dahl et al.'s confederate witnesses, in spite of the latter using scripts based on responses from real witnesses.

The next step is to extend this research to real police investigators. The investigators in Dahl et al.'s (2006) studies, as well as in the current research, were undergraduate students rather than actual investigators. It is possible that real investigators will not be as swayed by eyewitness evidence as the student-investigators in these studies. Real investigators have a lot more experience with witnesses and may be better able to tell when a witness is credible or not and adjust their beliefs accordingly. As well, investigators may be more aware of the fallibility of eyewitness evidence than undergraduate students who may not have general knowledge of eyewitness accuracy rates and may be more likely to take eyewitness accounts as fact compared to real investigators.

To date, very little research has examined the impact that eyewitnesses have on investigators. Together with Dahl et al. (2006), these two studies build on a unique new line of research attempting to determine how investigators use eyewitness evidence in making their decisions. This research is important because investigators are the gatekeepers of the justice system, conducting investigations that lead to people being charged for crimes they may or may not have committed. As cases of wrongful convictions accumulate, it is important to determine how these decisions are made to better inform the justice system.

References

Boccaccini, M. T., Gordon, T., & Brodsky, S. L. (2003). Effects of witness preparation on witness confidence and nervousness. *Journal of Forensic Psychology Practice, 3*(4), 39-51.

Brigham, J. C., & Bothwell, R. K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identifications. *Law & Human Behavior, 7*(1), 19-30.

Clark, S., Abbe, A., & Larson, R. (2006, November). Collaboration in associative recognition memory: Using recalled information to defend 'new' judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(6), 1266-1273.

Dahl, L., C., Lindsay, D. S., & Brimacombe, C. A. E. (2006). Investigating investigators: Examining witnesses' influence on investigators. *Law & Human Behavior, 30(6),* 707-733.

Kassin, S. M., & Neumann, K. (1997). On the power of confession evidence: An experimental test of the fundamental difference hypothesis. *Law & Human Behavior, 21*(5), 469-484.

Lindsay, D. S., Nilsen, E., & Read, J. D. (2000). Witnessing-condition heterogeneity and witnesses' versus investigators' confidence in the accuracy of witnesses' identification decisions. *Law and human behavior, 24*(6), 685-697.

Lindsay, R. C. L. (1994). Expectations of eyewitness performance: Jurors' verdicts do not follow from their beliefs. In D. F. Ross, & J. D. Read (Eds.), *Adult eyewitness testimony: Current trends and developments.* (pp. 362-384) Cambridge University Press.

Lindsay, R. C., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*(3), 556-564.

Lindsay, R. C., Wells, G. L., & O'Connor, F. J. (1989). Mock-juror belief of accurate and inaccurate eyewitnesses: A replication and extension. *Law & Human Behavior, 13*(3), 333-339.

Lindsay, R. C., Wells, G. L., & Rumpel, C. M. (1981). Can people detect eyewitness-identification accuracy within and across situations? *Journal of Applied Psychology, 66*(1), 79-89.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*(3), 203-220.

Wells, G. (2000). Eyewitness testimony. *Encyclopedia of Psychology, Vol. 3* (pp. 308-310). American Psychological Association.

Wells, G. L., & Bradfield, A. L. (1998). 'Good, you identified the suspect': Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83*(3), 360-376.

Wells, G., Ferguson, T., & Lindsay, R. (1981). The tractability of eyewitness confidence and its implications for triers of fact. *Journal of Applied Psychology*, *66*(6), 688-696.

Wells, G. L., Lindsay, R. C., & Tousignant, J. P. (1980). Effects of expert psychological advice on human performance in judging the validity of eyewitness testimony. *Law and human behavior, 4*(4), 275-285.

Wells, G. L., & Olson, E. A. (2002). Eyewitness identification: Information gain from incriminating and exonerating behaviors. *Journal of Experimental Psychology: Applied, 8*, 155-167.

Yuille, J., & Cutshall, J. (1989). Analysis of the statements of victims, witnesses and suspects. *Credibility assessment* (pp. 175-191). Kluwer Academic/Plenum Publishers.

Appendix

Equations for Posterior Probability that the Suspect is the Culprit given Eyewitness Identification Decision

$$p(\text{S is culprit}|\text{IDS}) = \frac{p(\text{IDS}|\text{S is culprit})p(\text{S is culprit})}{p(\text{IDS}|\text{S is culprit})p(\text{S is culprit}) + p(\text{IDS}|\text{S not culprit})p(\text{S not culprit})}$$

$$p(\text{S is culprit}|\text{IDfoil}) = \frac{p(\text{IDfoil}|\text{S is culprit})p(\text{S is culprit})}{p(\text{IDfoil}|\text{S is culprit})p(\text{S is culprit}) + p(\text{IDfoil}|\text{S not culprit})p(\text{S not culprit})}$$

$$p(\text{S is culprit}|\text{noID}) = \frac{p(\text{noID}|\text{S is culprit})p(\text{S is culprit})}{p(\text{noID}|\text{S is culprit})p(\text{S is culprit}) + p(\text{noID}|\text{S not culprit})p(\text{S not culprit})}$$

Note: S = Suspect, IDS = identification of suspect, IDfoil = identification of foil, noID = no identification made

Footnotes

1. Another fundamental difference between police officers and jurors is that whereas the latter are laypersons with respect to the criminal justice system, police investigators are professionals, often with substantial prior experience working with witnesses. We return to this issue in the General Discussion, but for present purposes it is moot because the research we report tested mock investigators.

2. The experimenter was not blind to condition which is not ideal as there is the possibility that the note-taker could influence the participants. However, information was recorded in the same manner for all participants by responding in one or two word answers to each of the questions asked (see Appendix F). Also, as indicated in the results section, there were no differences in any of the pre-identification measures to indicate that the experimenter affected any of the responses.

3. For example, looking only at target present lineups, a measure of accuracy can be obtained by calculating the number of identifications of the suspect divided by the number of identifications of the suspect plus foils, that is, given that the witness makes an identification, what is the probability that he or she identifies the culprit? Witnesses were significantly more likely to identify the culprit in the good view condition (56% of the time) than the poor view condition (20% of the time), $z = 1.72$, $p = .043$ (one-tailed). However, we are interested in accuracy when the suspect is present as well as when the suspect is not present as this is the case for concern in the real world. As witnesses were no more likely to make an accurate decision when they had a good view (56% of the time) than when they had a poor view (50% of the time) in target absent lineups, $z = .31$, $p = .378$, it was decided to collapse across viewing condition for all analyses.

4. Target presence is not included as a factor here to simplify the results as it was found that *Fs* < 1 for the main effect of target presence and for all interactions with target presence. In the real world, we wouldn't know if the culprit was present in the lineup or not; thus, we have collapsed across TP and TA conditions. However, details of these analyses are available from the first author.

5. See Wells and Bradfield (1998) for an explanation of why a MANOVA is the appropriate analysis here.

6. Target presence is not included as a factor here to simplify the results as it was found that *Fs* < 1.3 for the main effect of target presence and nearly all interactions where target presence was included as a factor. For the Target x ID Decision interaction, $F(1, 118) = 1.48$, $MSE = 506.95$, $p = .226$, partial $\eta^2 = .01$, and for the Time x View x Target x ID Decision interaction, $F(1, 118) = 2.12$, $MSE = 192.36$, $p = .148$, partial $\eta^2 = .02$.

7. We thank an anonymous reviewer for bring this contribution to our attention.

8. Statistical comparisons aren't made between Dahl et al. and Study 2 as different materials were used in the two studies. However, in Study 2, pre-identification probabilities were 63.66% (SD = 20.01%) and 65.35% (SD = 16.26%) for the good view and poor view conditions respectively.

Table 1

Frequency distribution of identification choices across target absent and target present lineups

	Target Absent		Target Present	
	Good View	Poor View	Good View	Poor View
Suspect ID	6	4	10	3
Foil ID	11	13	8	12
No ID	22	17	16	19

Table 2

Frequency distribution of identification choices across target absent and target present lineups

	Target Absent		Target Present	
	Good View	Poor View	Good View	Poor View
Suspect ID	8	8	25	9
Foil ID	11	19	7	17
No ID	35	18	9	14

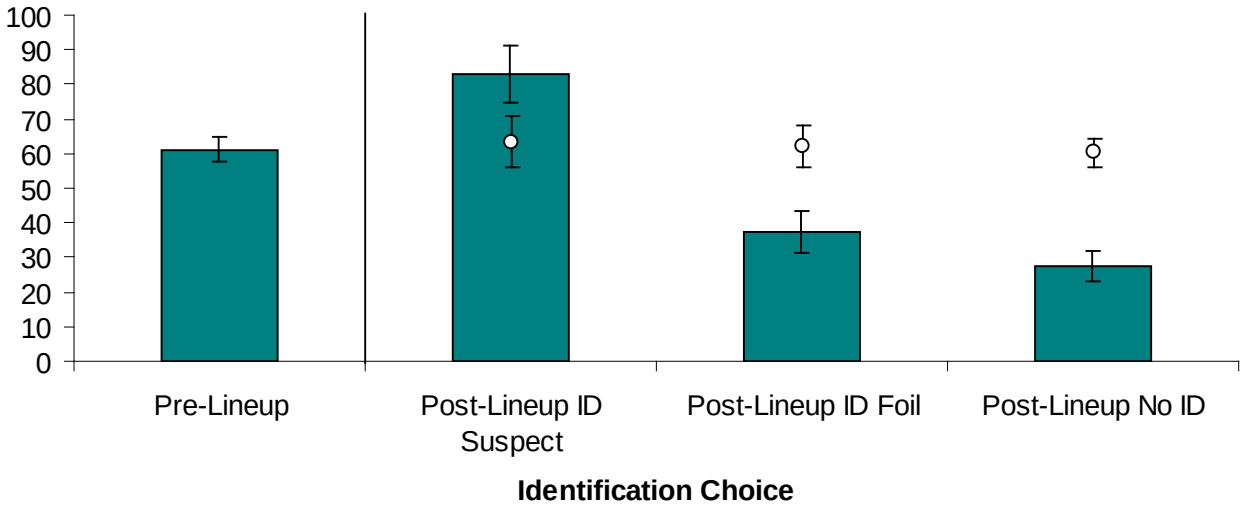
Figure Caption

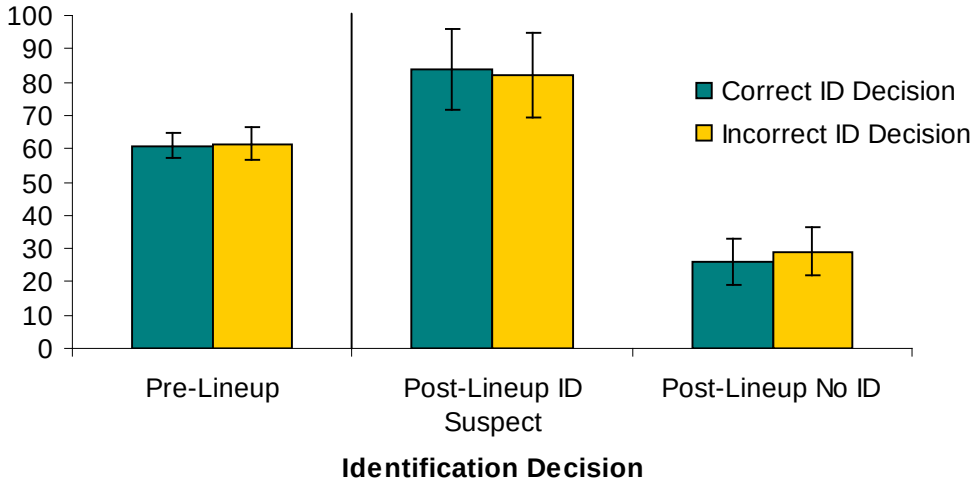
Figure 1. Investigators' percent probability suspect committed crime. The pre-lineup error bar represents a 95% within-subject confidence interval appropriate for comparisons between pre- and post-lineup. The post-lineup error bars represent 95% between-subjects confidence intervals appropriate for comparisons between post-lineup conditions (see Masson & Loftus, 2003). Circles indicate pre-lineup percent probability suspect committed crime plus or minus information gain from the lineup-task judgment with 95% confidence intervals.

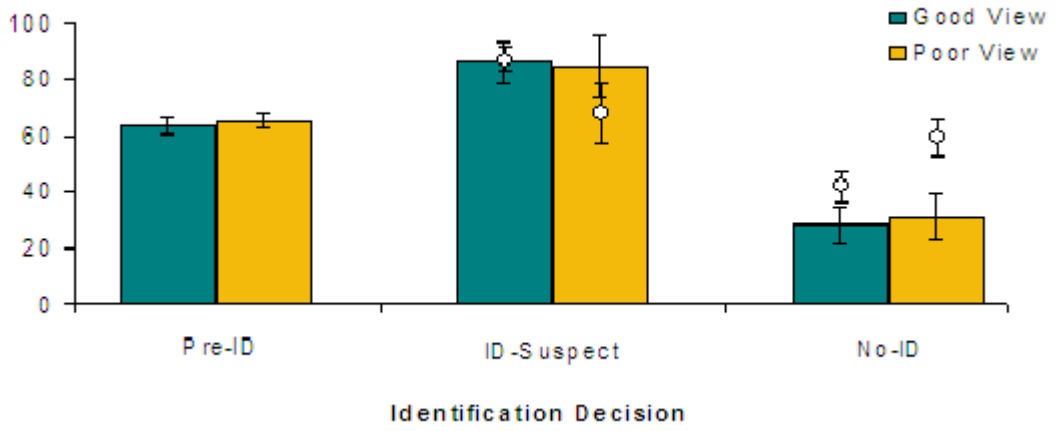
Figure 2. Investigators' percent probabilities suspect committed crime. Error bars represent 95% between-subjects confidence intervals.

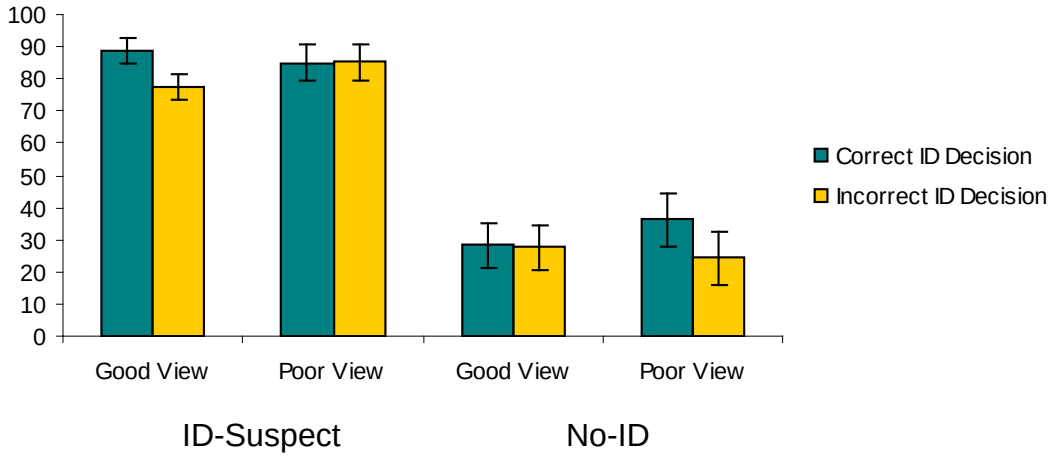
Figure 3. Investigators' percent probability suspect committed crime. The pre-lineup error bar represents a 95% within-subject confidence interval and is appropriate for comparisons between pre and post-lineup. The post-lineup error bars represent 95% between-subjects confidence intervals appropriate for comparisons between post-lineup conditions (see Masson & Loftus, 2003). Circles indicate pre-lineup percent probability suspect committed crime plus or minus information gain from the lineup-task judgment with 95% confidence intervals.

Figure 4. Investigators' percent probabilities suspect committed crime. Error bars represent 95% between-subjects confidence intervals.


100
90
80
70
60
50
40
30
20
10
0
Pre-Lineup
Post-Lineup ID Suspect
Post-Lineup ID Foil
Post-Lineup No ID
Identification Choice


100
90
80
70
60
50
40
30
20
10
0
Correct ID Decision
Incorrect ID Decision
Pre-Lineup
Post-Lineup ID Suspect
Post-Lineup No ID
Identification Decision


100
80
60
40
20
0
Good View
Poor View
Pre-ID
ID-Suspect
No-ID
Identification Decision


100
90
80
70
60
50
40
30
20
10
0
Good View
Poor View
Good View
Poor View
ID-Suspect
No-ID
Correct ID Decision
Incorrect ID Decision