Investigating Investigators:

Examining the Impact of Eyewitness Evidence on Student-Investigators

by

Melissa Ann Boyce

B.A.H., Queen's University, 2002
M.A., Queen's University, 2004

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Psychology

Investigating Investigators:

Examining the Impact of Eyewitness Evidence on Student-Investigators

By

Melissa Ann Boyce
B.A.H., Queen's University, 2002
M.A., Queen's University, 2004

Supervisory Committee

Dr. D. Stephen Lindsay, Co-Supervisor
(Department of Psychology)

Dr. C. A. Elizabeth Brimacombe, Co-Supervisor
(Department of Psychology)

Dr. Helena Kadlec, Departmental Member
(Department of Psychology)

Prof. Benjamin L. Berger, Outside Member
(Faculty of Law)

Supervisory Committee

Dr. D. Stephen Lindsay, Co-Supervisor
(Department of Psychology)

Dr. C. A. Elizabeth Brimacombe, Co-Supervisor
(Department of Psychology)

Dr. Helena Kadlec, Departmental Member
(Department of Psychology)

Prof. Benjamin L. Berger, Outside Member
(Faculty of Law)

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 in Studies 1 and 2 had been shown either a good or poor view of the perpetrator or in Study 3 viewed either a Caucasian or an Asian criminal, 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 generally did not differentiate between accurate and inaccurate witnesses.

Table of Contents

Supervisory Committee……..…ii

Abstract……iii

Table of Contents……..…iv

List of Tables……...vii

List of Figures……viii

Acknowledgments…….……..ix

Introduction…….……..…1

Rationale for Study 1…….…….…..…10

Methods…….……..….11

Participants…….…..11

Materials…….…….12

Procedure…….14

Results…….17

Discussion…….…..28

Rationale for Study 2……..30

Methods……..32

Participants…….32

Materials…….32

Procedure……....33

Results……33

Discussion……..47

Rationale for Study 3…….50

Methods…….…….51

Participants…….…….51

Materials……..…..…….52

Procedure……...….……54

Results……...….……54

Discussion……...…..66

General Discussion……...…..69

References……..78

Appendix A. Sample from Database: Information Contained in Suspect Database for John Gibbs in Study 1 ……....87

Appendix B. Investigator Pre-ID Questionnaire…….……..88

Appendix C. Investigator Post-ID Questionnaire…….……....89

Appendix D. Witness Pre-ID Questionnaire……91

Appendix E. Witness Post-ID Questionnaire…….……..92

Appendix F. Interview Instructions…….…….……93

Appendix G. Notes Page…….…….94

Appendix H. Instructions for Trying to Find a Suspect…….…….95

Appendix I. Lineup Checklist…….…….97

Appendix J. Equations for Posterior Probability that the Suspect is the Culprit Given Eyewitness Identification Decision……..…..…98

Appendix K. Sample from Database: Information Contained in Suspect Database for Jane Gibbs in Study 2……..…....99

Appendix L. Sample from Database: Information Contained in Suspect

Database for John Gibbs in Study 3……...100

Appendix M. Sample from Database: Information Contained in Suspect

Database for John Choi in Study 3…….……...101

List of Tables

Table 1. Frequency distribution of identification choices across target absent and target present lineups in Study 1……..……..…..……..……22

Table 2. Witnesses' mean responses across good and poor viewing conditions in Study 2…….……..38

Table 3. Frequency distribution of identification choices across target absent and target present lineups in Study 2…….……45

Table 4. Frequency distribution of identification choices across target absent and target present lineups in Study 3…….……59

List of Figures

Figure 1. Mean percentage ratings of each type of evidence…….20

Figure 2. Investigators' percent probability suspect committed crime……...24

Figure 3. Investigators' percent probabilities suspect committed crime……26

Figure 4. Mean percentage ratings of each type of evidence……37

Figure 5. Investigators' percent probability suspect committed crime……..41

Figure 6. Investigators' percent probabilities suspect committed crime…….……..44

Figure 7. Mean percentage ratings of each type of evidence…….……..58

Figure 8. Investigators' percent probability suspect committed crime…….61

Figure 9. Investigators' percent probabilities suspect committed crime…….……..64

Acknowledgments

I would like to acknowledge my supervisors, Steve Lindsay and Liz Brimacombe, for their help and support during the completion of my PhD. I would also like to acknowledge Leora Dahl and Carla Mclean for their input and ideas. Finally, I would like to thank my family for their encouragement every step of the way.

Introduction

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 typically cannot 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). The type of lineup used most often by police is the simultaneous lineup (Lindsay & Wells, 1985). In this lineup, members are viewed together. This may be done live or by using a photo spread. This lineup does not typically present a problem when the criminal is present. Although there are some exceptions, such as when the criminal's appearance has changed (Charman & Wells, 2007) or when the foils are too similar to the criminal (Wells, Rydell, & Seelau, 1993), witnesses tend to be quite adept at picking a guilty suspect from a simultaneous lineup (Lindsay & Wells, 1980). The problem arises when the criminal is not in the lineup, i.e., when police have arrested an innocent person. Even when the lineup is conducted properly, typical false identification rates (selection of the innocent suspect because of resemblance to the criminal) are about 25% to 45% (e.g., Malpass & Devine, 1981).

When biases are introduced, false identification rates rise substantially. Malpass and Devine (1981) tested the effects of an instruction bias (indicating that the criminal is in the

lineup) on identification accuracy and found a false identification rate of 78%. Studies examining the effects of clothing and foil biases have demonstrated significant increases in false identification rates as well (Lindsay, Wallbridge, & Drennan, 1987; Lindsay & Wells, 1980). A clothing bias exists when only the suspect is wearing an outfit similar to that worn by the criminal during the crime, or at least the criminal's clothing as described by the witness. A foil bias exists when the other members in the lineup do not fit the eyewitness' description to the same degree that the suspect does. Once again this causes the suspect to stand out. When these biases are combined, false identification rates skyrocket to as high as 89% (Lindsay, Lea, Nosworthy, & Fulford, 1991).

Yet eyewitnesses are not always in error (e.g., Leippe, Wells, & Ostrom, 1978; Yuille & Cutshall, 1989). Witnesses can be highly accurate in some situations. Some examples of factors that can lead to high accuracy rates for eyewitnesses include good lighting and vantage point, prolonged exposure time, absence of a weapon, and focused attention on the culprit (Haber & Haber, 2000, 2001). This raises the question of how well individuals involved in the criminal justice system can differentiate between accurate and inaccurate eyewitness evidence. The focus of the current research is how those in the role of investigating officer evaluate eyewitnesses' identification evidence.

Prior research has indicated that mock jurors consistently overestimate eyewitnesses' accuracy regardless of the gender, age, education, and real-world jury experience of the mock juror 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). Lindsay et al. (1981) varied the viewing conditions that witnesses had in order to manipulate witness accuracy. They found that although mock jurors

appeared sensitive to witnesses' viewing conditions, decreasing their judgments of witness accuracy as the viewing conditions worsened, this effect mainly occurred only if witnesses were low in confidence. In addition, mock jurors consistently believed that witnesses were more likely to be accurate than they were even when viewing conditions led to relatively high levels of accuracy.

Lindsay et al. (1989) had lawyers cross-examine witnesses to a staged crime in a courthouse setting. They then showed these cross-examinations to mock jurors to see whether mock jurors would better be able to judge eyewitness accuracy under these more realistic conditions. However, they found that even when allowed to watch witness cross-examinations, mock jurors still overestimated the accuracy of eyewitness evidence and did not distinguish between accurate and inaccurate eyewitnesses.

There is no published research that has specifically compared investigators to jurors; however, it is possible that investigators may differ from jurors in these regards for several reasons. In order to illustrate the ways in which the decision making of jurors and investigators may differ, it is helpful to outline the tasks of each.

The Juror's Task:

The ultimate function of every juror is to make a judgment regarding whether an accused person is legally guilty of the crime for which he or she has been tried. Thus, the juror's job is not to decide whether he or she feels the person is guilty or not but whether the evidence is "sufficiently convincing to establish a very strong probability of guilt" (Gelfand & Soloman, 1973). To perform this function, each juror must consider all of the information presented at trial and determine based on this information if it appears likely that the defendant committed the crime subject to the standard of reasonable doubt.

Legally, there has been some controversy over how much detail (if any) needs to be provided to help jurors understand this concept (Dane, 1985; R. v. Lifchus, 1997). However, in Canada, several guidelines have been set in place to help courts instruct jurors on the meaning of reasonable doubt (R. v. Lifchus, 1997). First, the rationale for requiring proof beyond a reasonable doubt is based on the presumption of innocence, which is a fundamental principle of the criminal justice system in Canada and many other countries. Because of this, the burden of providing proof beyond a reasonable doubt falls on the prosecution, not on the accused. In addition, reasonable doubt should not be based on "sympathy or prejudice," but on "reason and common sense" based on the evidence presented at trial (or lack thereof). Finally, the proof does not need to be absolute, but must be more conclusive than showing probable guilt. If the evidence against an accused is not of sufficient strength to provide proof of guilt beyond a reasonable doubt according to these standards, then the juror is to err on the side of caution and return a verdict of not guilty.

From a psychological perspective, Hastie (1993) outlined a series of subtasks that each juror must perform in order to reach a verdict decision. First, being exposed to information during the course of a trial is not enough. Jurors are intended to pay attention to, remember, and appropriately interpret the meaning of all of the information presented to them during the trial, which will include testimony from witnesses, evidence or exhibits presented, and arguments from lawyers.

At the end of the trial, the judge provides the jury with procedural and verdict instructions to help them appropriately interpret and apply this information, and jurors are meant to understand and remember these instructions as well. Procedural instructions include instructions regarding the presumption of innocence, determination of the facts, admissibility, credibility,

reasonable inference, and reasonable doubt (Hastie, 1995). Verdict instructions tell jurors about the different verdicts possible in the case (e.g., guilty of first degree murder, guilty of second degree murder, guilty of manslaughter, not guilty, etc.) and the features constituting each verdict category which must be satisfied to find the accused guilty.

At this stage, jurors are supposed to consider only the evidence that has been deemed admissible, arguably a difficult task, particularly when jurors have been instructed to ignore inadmissible evidence (see Wolf & Montgomery, 1977). They must then evaluate the credibility and implications of this evidence to produce an account of the events of the crime. Additional evidence may be sought by jurors through inference or revisiting the evidence presented to come up with a plausible sequence of events. Finally, a predeliberation verdict is reached based on the strength of the case and taking into account the judge's procedural and verdict instructions.

The Investigator's Task:

The ultimate goal of an investigator in a criminal case is to find the person or persons responsible for committing a particular crime. To achieve this goal, investigators must actively seek out and synthesize information relevant to the case to determine and build a case against a suspect. The legal standard that must be met for an arrest to be made is that the investigator must have "reasonable and probable grounds" to believe that a suspect committed a crime (R v. Storrey, 1990). These grounds must be more than subjective but *are* based on a common-sense approach in that "a reasonable person placed in the position of the officer" would also feel that reasonable and probable grounds exist given the same circumstances (R. v. Storrey, 1990).

From a psychological perspective, the task of the investigator also involves several subtasks. Given that investigators of a crime are typically absent during the commission of that crime, their first objective when they begin their investigation is to seek out as much information

as possible to determine what happened. This is done by examining physical evidence, speaking with forensic experts, and collecting statements from witnesses, victims, potential suspects, and the general public (Yarmey, 2001). Unlike jurors, who are exposed only to information that has already been deemed relevant to the case at hand, investigators must sort through all of the information to which they have been exposed, deciding what information is important and relevant. Investigators are meant to consider all of this information objectively in order to develop an account of what happened during the crime and who is most likely to have committed the crime.

In many cases, the information received will lead investigators to consider one or more potential suspects, each of whom is investigated further. It is at this stage that innocent suspects should be ruled out based on further investigation as investigators attempt to narrow the list down to one suspect. At this point, investigators may make an arrest if they have enough evidence against that suspect to have objective grounds to believe the suspect committed the crime according to the principles of "reasonable and probable grounds." If not, they may attempt to build a case against that suspect, finally making an arrest if enough evidence is accrued that the legal standard has been met.

Comparison of Juror and Investigator Decision-Making

There are clearly important differences between the responsibilities of the investigator compared to those of the juror. Although speculative, 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. On the other hand, there is some evidence that investigators may be more likely to believe a suspect is guilty (i.e., exhibit a guilty bias) than jurors who are instructed to

adhere to the principle of "beyond a reasonable doubt" when deciding whether to vote guilty or not (e.g., Meissner & Kassin, 2002).

In addition, 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, play an active role in building a case by establishing a suspect, collecting evidence, interviewing witnesses, and finally deciding whether or not to arrest that suspect. Natter and Berry (2005) found that people who took an active role in gathering information understood and applied risk information more accurately.

Another difference relates to the type of witnesses jurors see in comparison with investigators. It seems likely that only witnesses who are reasonably confident will end up testifying at trial, even though eyewitness confidence may not be the best indicator of accuracy (Leippe, 1980). Research indicates that post-identification feedback that an eyewitness was correct in his or her identification can significantly inflate eyewitness confidence (e.g., Semmler, Brewer, & Wells, 2004; Wells, Olson, & Charman, 2003). Even if the eyewitness was not told that he or she had identified the suspect at the time of the identification (as recommended by the National Institute of Justice, 1999), the witness will know that he or she has identified the suspect when the suspect is charged with the crime. In addition, lawyers often rehearse witnesses' testimony with them prior to their appearance in court and research shows that witnesses who have practiced giving their statements show increased confidence (Boccaccini, Gordon, & Brodsky, 2003). If only witnesses who are reasonably confident in the first place are asked to testify at trial, then once combined with these factors that also increase confidence it is likely that most witnesses who testify in court show an overwhelming display of confidence.

Confidence has consistently been shown to have a substantial influence on whether a witness is believed (e.g., Cutler, Penrod, & Dexter, 1990; Cutler, Penrod, & Stuve, 1988).

There may be other differences between witnesses who make it to trial and those who don't. For example, by the time witnesses appear in court, delay and other various psychosocial processes may attenuate differences between accurate and inaccurate witnesses (e.g., Wells, Ferguson, & Lindsay, 1981). Jurors only see witnesses who make it to trial whereas investigators see all witnesses who are interviewed. 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]

Research on Investigators

Almost all research examining belief of eyewitnesses 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.

Studies have focused on the impact that investigators can have on eyewitnesses, for example, through their questioning style (Poole & White, 1991), by suggesting information to the witness (Wright, Self, & Justice, 2000), or through lineup administrators' expectations (Garrioch & Brimacombe, 2001). However, very few studies have focused on the effects that the eyewitness has on the investigator. There are only two published studies that have focused on

[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. This issue is revisited in the General Discussion, but for present purposes it is moot because this research tested mock investigators.

the effects of eyewitnesses on investigators (Dahl, Lindsay, & Brimacombe, 2006; Lindsay, Nilsen, & Read, 2000).

Lindsay et al. (2000) looked at mock investigators' ability to discriminate between accurate and inaccurate witnesses. Lindsay et al. randomly assigned participants to be witnesses or play the role of an investigator. Witnesses watched a video of a crime, and viewing conditions were manipulated so that witnesses were either exposed to a good or a poor view of the crime. Investigators then administered a photo lineup to witnesses and asked them a series of questions with the aim of assessing the witnesses' accuracy. Lindsey et al. found that investigators distinguished between accurate and inaccurate witnesses to some extent as they showed more confidence in witnesses who had a good view than witnesses who had a poor view. However, they discriminated between accurate and inaccurate witnesses more poorly than the witnesses themselves. Investigators also had a tendency to accept eyewitness identification decisions, being significantly more likely to judge correct identifications as accurate than to judge incorrect identifications as inaccurate.

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 database provided a physical description of each potential suspect, along with information pertaining to prior arrest record, current residence and occupation, and alibi.

The investigator rated the probability that his or 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 all members of the lineup.

Rationale for Study 1

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 et al., 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.

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 did not 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, presumably 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 by witnesses who had a poor view.

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. Appendix A includes this suspect's information.

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 his 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 photo was of the culprit and for TA lineups the photo was of 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. Copies of these two questionnaires are included in Appendices B and C.

In addition, witnesses filled out one questionnaire prior to viewing the lineup and another questionnaire immediately afterward. The questionnaires were adapted from Bradfield and Wells (1998) and asked witnesses 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. The questionnaires are included in Appendices D and E.

Procedure

Each pair of participants 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). These instructions are included in Appendix F.

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] (see Appendix G). 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

[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 G). 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.

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. The complete instructions are included in Appendix H. 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 with the names of each of the potential suspects and encouraged 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. Three (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. The instructions are included in Appendix I.

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 indicate who they had selected as their suspect, to rate the likelihood that their suspect was guilty on a scale from 0% to 100%, to indicate whether they would arrest the suspect at this point in the investigation, and to report what factors had contributed to their selection of their suspect. Target presence was included as a factor in the majority of the pre-identification analyses to ensure that the experimenter had not affected any of these measures as she was not blind to condition.

Selection of suspect. Investigators first indicated which suspect they had chosen to put in the lineup. John Gibbs was selected by 51.1% (n = 72) of participants. Another 21.3% (n = 30) chose William White, 14.9% (n = 21) chose Roberto Romero, and 5.7% (n = 8) chose Francis Beauchamp. The rest of the choices were scattered between the other suspects, with all suspects but one (Juan Martinez) being chosen by at least one investigator.

Investigator's pre-identification probabilities suspect committed crime. Overall, investigators indicated that there was a 61.22% (SD = 18.14) chance that their suspect was guilty prior to conducting the lineup. 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, a 2 (Target: Present, Absent) x 3 (ID

Decision: ID Suspect, ID Foil, No ID) ANOVA was conducted on investigators' pre-ID probabilities of guilt. However, pre-identification probabilities did not differ as a function of witness presence or identification decision, nor was the interaction significant (all *F*s < 1 except for ID Decision: $F(2, 132) = 1.32$, $MSE = 334.08$, $p = .271$, partial $\eta^2 = .02$).

There were pre-identification differences, however, in ratings of the likelihood the suspect committed the crime depending on which suspect was chosen by the investigator, $F(8, 129) = 3.95$, $MSE = 280.82$, $p < .001$, partial $\eta^2 = .20$. Pair-wise analyses using the Bonferroni correction (excluding Tom Rabin and Hans Lindbolm as each was selected only once) indicated that investigators were significantly more confident in their suspect's guilt if they'd chosen John Gibbs ($M = 67.16$, $SD = 14.04$) than if they'd chosen William White ($M = 55.5$, $SD = 21.47$, $t(98) = 2.74$, $p = .010$, $d = .64$[3]), Hank Ellis ($M = 35.0$, $SD = 13.23$, $t(71) = 4.11$, $p < .001$, $d = 2.36$), or Nigel Ames ($M = 23.5$, $SD = 23.33$, $t(70) = 2.63$, $p = .010$, $d = 2.27$).

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* <

[3] Cohen's *d* is provided whenever a t-test on means is reported. Cohen's *d* is equal to the difference between 2 means divided by the pooled standard deviation for those means, where an effect size of .20, .50, and .80 correspond to small, medium, and large effect sizes respectively (Cohen, 1992).

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 η^2 = .06. The causal implications of this finding are ambiguous, given that arrest decision was not manipulated.

To determine whether investigators' decision to charge depended on which suspect they had chosen, the top three suspects chosen by investigators (John Gibbs, William White, and Roberto Romero) were kept separate and all other suspects chosen grouped into the category of "other" in order to conduct chi-square analyses. As the database was designed to favor John Gibbs as the suspect, there was a possibility that investigators who had chosen John Gibbs would be more confident than investigators who had chosen someone else. However, the decision to arrest did not depend on which suspect investigators had chosen, $X^2(3) = 1.20$, p = .753, V = .09. This was true when those who chose John Gibbs were compared to those who chose anyone else as well, $X^2(1) = .06$, p = .803, V = .02. The suspect selected (John Gibbs, William White, Roberto Romero, Others or John Gibbs vs. Others) also did not significantly affect investigators' confidence in their decision to charge or not charge the suspect (Fs < 1).

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 (*F*s < 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. As can be seen from Figure 1, physical description was rated as significantly more important than prior criminal record ($t(139) = 12.09$, $p < .001$, $d = 1.38$), fingerprints ($t(139) = 8.41$, $p < .001$, $d = 1.09$), alibi ($t(139) = 10.33$, $p < .001$, $d = 1.24$), and other evidence ($t(139) = 17.46$, $p < .001$, $d = 2.23$). Prior criminal record was rated as significantly more important than other evidence, $t(139) = 6.86$, $p < .001$, $d = 0.80$. Fingerprints were rated as significantly more important than prior criminal record and other evidence ($t(139) = 3.49$, $p = .001$, $d = 0.31$ and $t(139) = 9.37$, $p < .001$, $d = 1.15$ respectively). Finally, alibi was rated as significantly more important than other evidence, $t(139) = 8.70$, $p < .001$, $d = 1.06$.

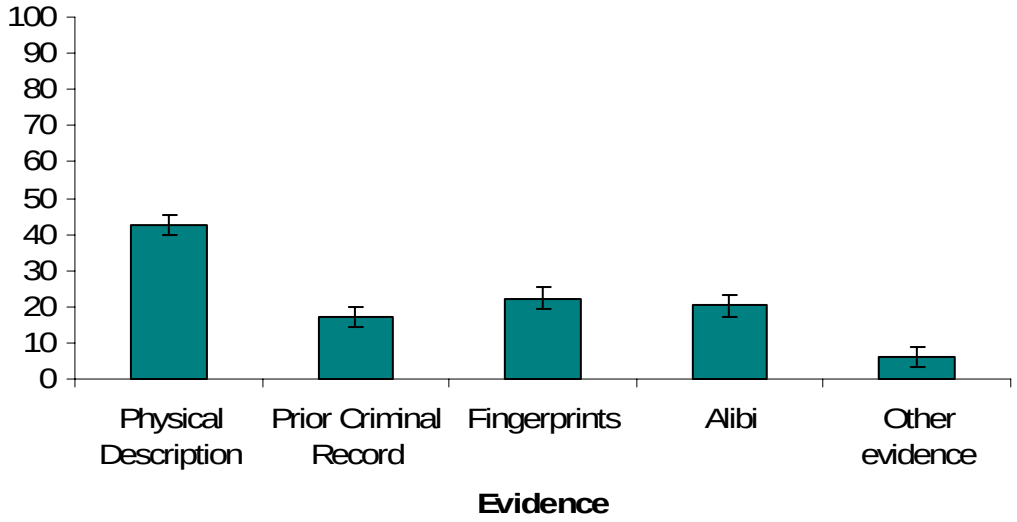

Figure 1. Mean percentage ratings of each type of evidence. The error bars represent the 95% within-subject confidence interval, appropriate for comparing the various kinds of evidence (see Masson & Loftus, 2003).

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[4], 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, $z = .22$, $p = .823$.

[4] 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, I am 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. I return to the viewing condition issue in Study 2.

Table 1. *Frequency distribution of identification choices across target absent and target present lineups in Study 1.*

	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

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 diagnosticity of the judgment their witness had given 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[5] was

[5] Target presence is not included as a factor here to simplify the results as *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, I have collapsed across TP and TA conditions.

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 ($M = 83.04$, $SD = 14.48$), $t(22) = -8.09$, $p < .001$, $d = -1.61$. If the lineup was rejected, investigators' guilt probabilities went down significantly from pre-lineup ($M = 61.05$, $SD = 17.04$) to post-lineup ($M = 27.43$, $SD = 19.78$), $t(71) = 13.03$, $p < .001$, $d = 1.82$. Finally, when a foil was identified, guilt probabilities dropped significantly from pre-lineup ($M = 64.07$, $SD = 19.65$) to post-lineup ($M = 37.26$, $SD = 21.14$), $t(42) = 7.41$, $p < .001$, $d = 1.31$.

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 and hence, the investigators' perceived probability that the suspect was the culprit tended to decline from pre to post-ID. 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 ($M = 37.26$, $SD = 21.14$) ($t(65) = 10.43$, $p < .001$, $d = 2.53$) or rejected the lineup ($M = 27.43$, $SD = 19.78$) ($t(95) = 14.65$, $p < .001$, $d = 3.21$). 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, $t(116) = 2.50$, $p = .014$, $d = 0.48$. These results can be seen in Figure 2.

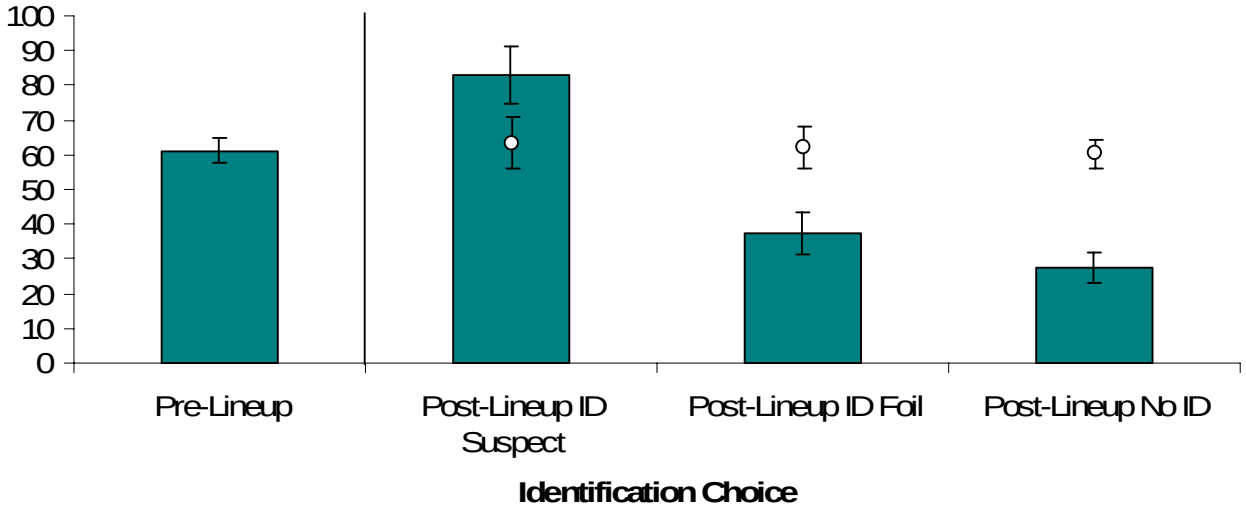

Figure 2. Investigators' percent probability suspect committed crime. The pre-lineup error bar represents a 95% within-subject confidence interval appropriate for comparisons between pre- and post-lineup. The post-lineup error bars represent 95% between-subjects confidence intervals appropriate for comparisons 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 around the individual cell means.

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 Appendix J. 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[6].

Circles are included on Figure 2 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 ($M = 63.42$, $SD = 17.11$), $t(22) = 6.22$, $p < .001$, $d = 1.23$. Conversely, when a foil was identified or the lineup was rejected, investigators returned significantly lower probabilities of guilt (Foil ID: *M*

[6] Investigators' pre-ID probabilities were chosen as the prior probability that the culprit was in the lineup rather than the actual probability based on the rationale that there was no way for investigators to be aware of the actual probability that the culprit was in the lineup as that probability was set by the experimenter and the actual probability (.48) is unlikely to reflect real-world probabilities that a lineup contains the culprit. Thus, it made more sense to use investigators' pre-ID probabilities as these values reflected their own beliefs regarding the likelihood that the suspect was the culprit which could then easily be connected to the calculated posterior probabilities of the likelihood that the suspect was guilty based on each investigator's initial beliefs.

= 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, *d* = -1.20 and *t*(71) = -12.70, *p* < .001, *d* = -1.77 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 3, 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, *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, *d* = -0.22. This is not surprising given how close to chance witnesses were; many correct responses were likely to be guesses.

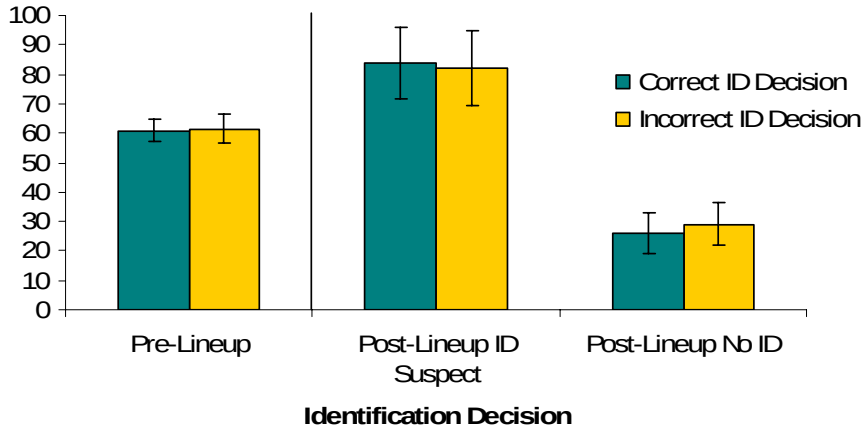

Figure 3. Investigators' percent probabilities suspect committed crime. Error bars represent 95% between-subjects confidence intervals, appropriate for comparing correct and incorrect ID decisions (see Masson & Loftus, 2003).

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) 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.

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, 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.

For ratings of how important physical description, prior criminal record, fingerprints, alibi, and other evidence were to investigators in choosing their suspect, physical description was rated as significantly more important than all other types of evidence. Conversely, other evidence was rated as being significantly less important than all other types of evidence.

It may be somewhat surprising at first that physical description was rated as being significantly more important than all other types of evidence, especially considering that fingerprint evidence was included for some suspects, as past research indicates that fingerprint evidence can be at least as compelling as eyewitness evidence (e.g., Bregman & McAllister, 1987). However, the fingerprint evidence was given in terms of a percent probability that the print taken from the crime scene matched that of the suspect and the highest probability that was given was 70% for John Gibbs. It is possible that people didn't give as much weight to a partial match and so didn't factor it into their decisions as much as the physical description they had obtained from the eyewitness, who was presumably credible having seen the crime. This also fits with past research indicating that eyewitness evidence is very compelling (Lindsay, 1994).

Comparison to Dahl et al. (2006)

It is worth providing a brief comparison of the results of the current study to the findings of Dahl et al. (2006). Both studies used the same materials and followed the same general paradigm; the important difference between the two studies is that Dahl et al. used confederate witnesses who provided scripted descriptions to investigators whereas the current research used real witnesses. Dahl et al.'s main findings were replicated in Study 1. Investigators were

dramatically affected by eyewitness identification decisions in Dahl et al.'s research as well, indicating that their results generalize to the impact of real witnesses on investigators.

The main difference between the findings of Study 1 and Dahl et al.'s (2006) research relates to suspect selections where there was substantially more variability in Study 1 than in Dahl et al.'s research using confederate witnesses. Whereas all investigators in their research chose John Gibbs as the suspect, only half did in Study 1. This difference is believed to be because the descriptions given by the witnesses in Study 1 were genuine, rather than scripted as in Dahl et al.'s research. 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. The witnesses in Dahl et al.'s pilot studies provided written descriptions of the criminal rather than verbal descriptions. Possibly they may have provided more detailed responses because they had time to think about their answers and to read them over to see if they forgot to mention anything.

If all witnesses in Study 1 had given detailed and accurate descriptions of the criminal they had seen, then most or all investigators would have selected 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.

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, the culprit looks different in the lineup photo than in the video. 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 Study 2, 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 typically 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 undergraduate 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. The information provided for Jane Gibbs is available in Appendix K.

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 (see Appendices B, C, & D).

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 indicate who they had selected as their suspect, to rate the likelihood that their suspect was guilty on a scale from 0% to 100%, to indicate whether they would arrest the suspect at this point in the investigation, and to report what factors had contributed to their selection of their suspect. Target presence was included as a factor in the majority of the pre-identification

analyses to ensure that the experimenter had not affected any of these measures as she was not blind to condition.

Selection of suspect. There were no differences between the good and poor conditions in terms of suspect chosen. In the poor view condition, 84.7% of participants chose Jane Gibbs (*n* = 72) and 15.3% chose other suspects (*n* = 13) whereas in the good view condition, 84.2% of participants chose Jane Gibbs (*n* = 80) and 15.8% chose other suspects (*n* = 15), $X^2(1) = .01$, *p* = .920, *V* = .01. Overall, 84.4% of investigators chose Jane Gibbs as the suspect (*n* = 152). The rest of the choices were scattered between the other suspects, although three suspects were not chosen by anyone (Amanda Hoffman, Kelly Vinson, and Maria Martinez).

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) x 3 (ID Decision: ID Suspect, ID Foil, No ID) 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 ($F < 1$), target presence ($F(2, 168) = 1.63$, $MSE = 326.80$, $p = .204$, partial $\eta^2 = .01$), or the witness's subsequent identification decision ($F < 1$). There was a significant View x ID Decision interaction ($F(2, 168) = 3.32$, $MSE = 326.80$, $p = .038$, partial $\eta^2 = .04$). Investigators provided significantly lower pre-identification probabilities of guilt when a foil was subsequently identified in the good view condition ($M = 55.39$, $SD = 24.79$) than in the poor view condition ($M = 69.22$, $SD = 14.43$), $t(52) = 2.59$, $p = .012$, $d = -.68$ (for Suspect ID: Good View: $M = 66.76$, $SD = 18.53$, Poor View: $M = 63.00$, $SD = 15.79$, $t(48) = .71$, $p = .479$, $d = .22$, for No ID: Good View: $M = 64.72$, $SD = 18.42$, Poor View: $M = 62.25$, $SD = 17.94$, $t(74) = .58$, $p = .561$, $d = .14$). However, this finding is believed to be a type I error, as there was more variability in investigators' pre-identification

probabilities that the suspect was guilty in the good view condition when a foil was subsequently identified in comparison to the poor view condition, likely as result of the fact that only 18 witnesses in the good view condition identified a foil. All other interactions did not reach significance (all *F*s < 1, except for View x Target x ID Decision: $F(2, 168) = 1.61$, $MSE = 326.80$, $p = .203$, partial $\eta^2 = .02$).

The investigators' suspect choices were separated into Jane Gibbs versus all others and those who chose Jane Gibbs as their suspect ($M = 67.12$, $SD = 16.69$) provided significantly higher pre-identification probabilities that she was guilty compared to those who chose someone else as their suspect ($M = 50.00$, $SD = 20.14$), $F(1, 178) = 23.27$, $MSE = 297.86$, $p < .001$, partial $\eta^2 = .12$. 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 replicate Study 1, other than the finding that in Study 1, rather than merely approaching significance, investigators were significantly more confident in their arrest decision if they chose to charge the suspect.

To determine whether investigators' decision to charge depended on which suspect they had chosen, Jane Gibbs was separated from the other suspects chosen who were grouped into the category of "other" in order to conduct chi-square analyses. The decision to arrest did not depend on which suspect investigators had chosen, $X^2(1) = 2.52$, $p = .112$, $V = .12$. The suspect chosen may have had more subtle effects on investigators however, as investigators were significantly more confident in their arrest decision when they had chosen Jane Gibbs as their suspect ($M = 6.60$, $SD = 1.56$) than when they had selected someone else ($M = 5.64$, $SD = 1.85$), $t(178) = 2.90$, $p = .004$, $d = .56$. Recall that the database was designed to favor Jane Gibbs as the most likely 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, 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. As can be seen from Figure 4, physical description was rated as significantly more important than fingerprints ($t(177) = 7.33$, $p < .001$, $d = 0.95$), alibi ($t(176) = 14.31$, $p < .001$, $d = 1.71$), prior criminal record ($t(176) = 14.77$, $p < .001$, $d = 1.65$), and other evidence ($t(172) = 25.92$, $p < .001$, $d = 2.80$). Fingerprints were rated as significantly more

important than alibi, prior criminal record, and other evidence ($t(176) = 6.25$, $p < .001$, $d = 0.64$, $t(176) = 5.49$, $p < .001$, $d = 0.63$, and $t(172) = 15.99$, $p < .001$, $d = 1.77$ respectively). Finally, alibi and prior criminal record were rated as significantly more important than other evidence, $t(173) = 15.01$, $p < .001$, $d = 1.59$ and $t(173) = 13.03$, $p < .001$, $d = 1.34$. 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.

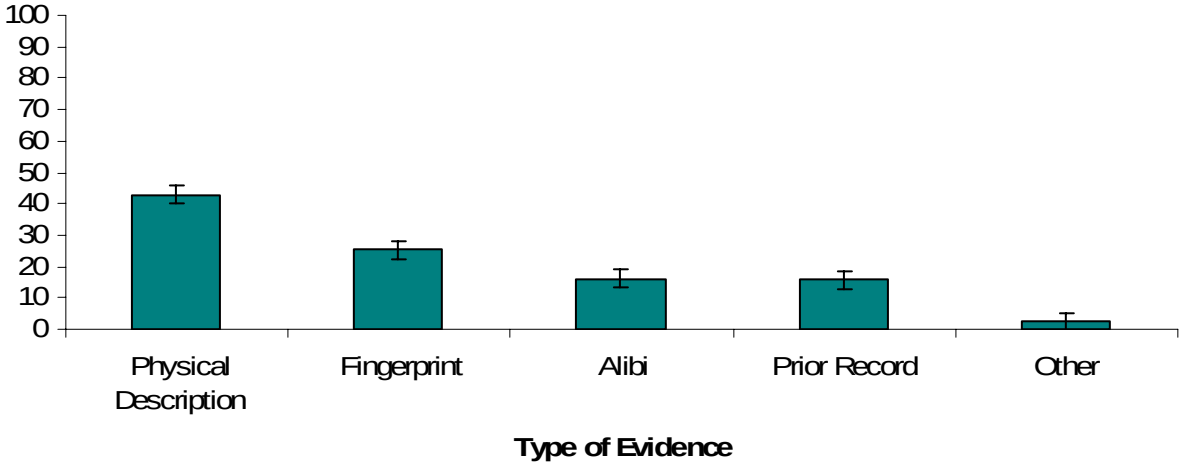

Figure 4. Mean percentage ratings of each type of evidence. The error bars represent the 95% within-subject confidence interval, appropriate for comparing the various kinds of evidence (see Masson & Loftus, 2003).

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.

Table 2. *Frequency distribution of identification choices across target absent and target present lineups in Study 2.*

	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

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. Therefore these data have been excluded from subsequent analyses.

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 diagnosticity of the judgment their witness had given 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.[7] 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 ($M = 85.60$, $SD = 11.62$), $t(49) = -9.56$, $p < .001$, $d = -1.35$. If the lineup was

[7] Target presence is not included as a factor here to simplify the results as $Fs < 1.3$ 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 $\eta^2 = .02$.

rejected, investigators' guilt probabilities dropped significantly from pre-lineup (M = 63.68, SD = 18.14) to post-lineup (M = 29.30, SD = 22.92), $t(75) = 13.53$, $p < .001$, $d = 1.66$.

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 and as a result, the investigators' perceived probability that the suspect was the culprit tended to decline from pre to post-ID. 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 ($t(124) = -16.05$, $p < .001$, $d = 3.10$) but not pre-identification probabilities ($F < 1$). 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 5. These results are similar to Study 1.

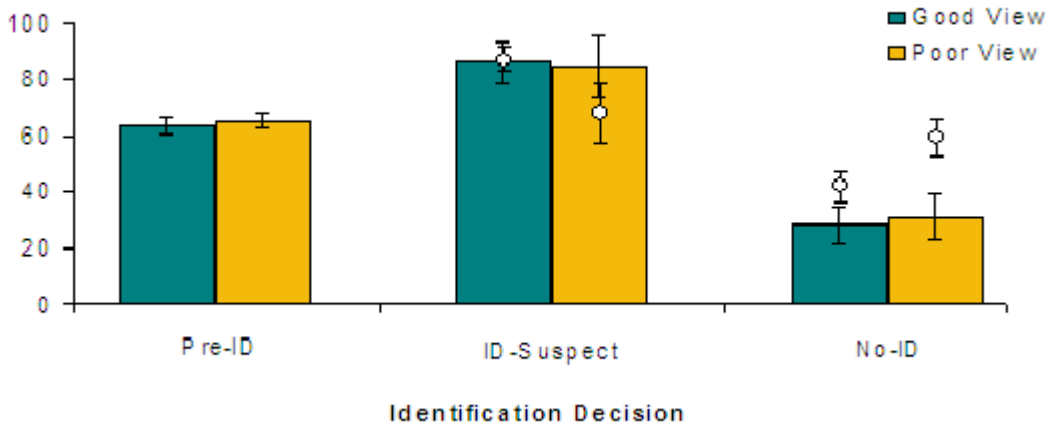

Figure 5. Investigators' percent probability suspect committed crime. The pre-lineup error bar represents a 95% within-subject confidence interval and is appropriate for comparisons between pre and post-lineup. The post-lineup error bars represent 95% between-subjects confidence intervals appropriate 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 around the individual cell means.

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 from 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 from 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 J). Circles on Figure 5 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), $t(16) = 5.65$, $p < .001$, $d = 1.17$ and $t(31) = -8.31$, $p < .001$, $d = -1.33$ 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), $t(43) = -3.70$, $p = .001$, $d = -.69$. 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), $t(31) = -.59$, $p = .558$, $d = -.11$. As discussed in greater detail in the General Discussion, I speculate that if conditions led eyewitnesses to be extremely accurate, such that information gain was very high, investigators would be under-influenced by them.

Did investigators 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 as a function of whether the identification was correct or incorrect for both the good (correct: M = 28.23, SD = 23.39, incorrect: M = 27.78, SD = 17.34) and poor viewing conditions (correct: M = 36.11, SD = 25.06, incorrect: M = 24.11, SD = 22.31, $t(42) = -.054$, p = .957, $d = 0.02$ and $t(30) = -1.40$, $p = .173$, $d = 0.17$ 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, $t(15) = .02$, $p = .986$, $d = -0.01$), which is similar to Study 1, in the good viewing condition 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, $t(31) = -3.00$, $p = .005$, d = 1.21). These results can be seen in Figure 6.

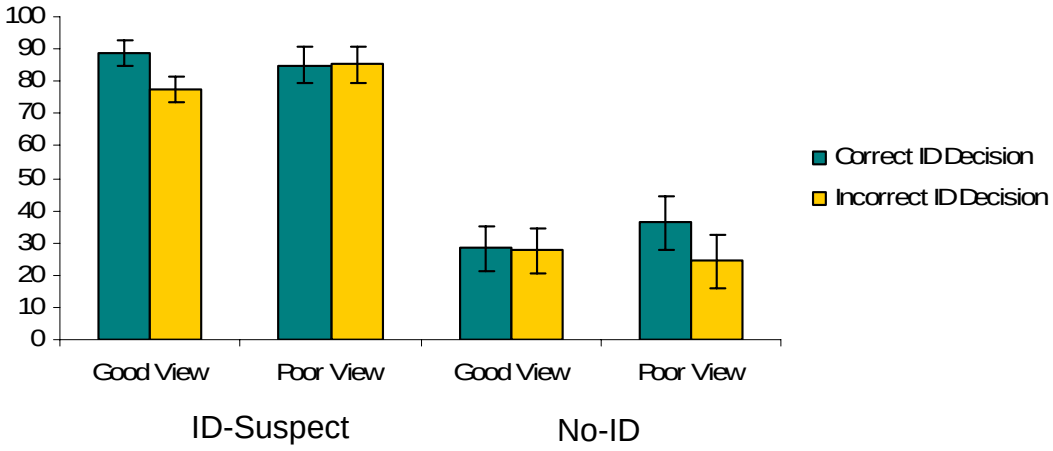

Figure 6. Investigators' percent probabilities suspect committed crime. Error bars represent 95% between-subjects confidence intervals, appropriate for comparing correct and incorrect ID decisions (see Masson & Loftus, 2003).

Eyewitness Questionnaires

Eyewitnesses also filled out a questionnaire on which they made a number of ratings about their bases for making their lineup response. 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, $F(12, 73) = 2.80$, $p = .003$. The means and standard deviations for each measure as a function of view can be seen in Table 3. 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 $\eta^2 = .09$), being able to make out the criminal's face better ($F(1, 84) = 18.37$, *MSE* = 1.77, $p < .001$, partial $\eta^2 = .18$), 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 $\eta^2 = .08$).

Table 3. *Witnesses' mean responses across good and poor viewing conditions in Study 2. Standard deviations are provided in brackets. Asterisks indicate values that differ at the .05 level.*

	Poor View	Good View
How many seconds would you estimate that the criminal's face was in view?	56.44 (64.18)	113.06 (104.17) *
How well were you able to make out the criminal's face from the video?	3.55 (1.17)	4.73 (1.30) *
What would you estimate was the closest distance between the camera-eye view and the criminal's face?	8.87 (2.54)	6.51 (2.54) *
How much attention were you paying to the criminal's face while viewing the video?	3.66 (1.21)	4.11 (1.13)
How confident are you that you could ID the criminal (if he was present) from a police lineup?	4.26 (1.37)	5.17 (1.44) *

	Poor View	Good View
At the time that you made your ID decision, how certain were you that the person you saw in the video was/was not in the photo lineup?	3.85 (1.46)	4.29 (1.54)
Witness estimate of probability ID decision was correct	54.98 (21.72)	64.85 (21.55) *
After you were first shown the photos, how long do you estimate it took you to make your decision?	4.18 (1.54)	3.69 (1.49)
How willing would you be to testify in court as to your ID decision?	3.04 (1.49)	3.88 (1.60) *
If another eyewitness had about the same view of the criminal, do you think an ID decision by this eyewitness ought to be trusted?	4.02 (1.35)	4.25 (1.39)
To what extent do you feel that you had a good basis (enough information) to make an identification?	3.86 (1.15)	4.47 (1.45) *

Was eyewitness confidence correlated with other variables?

Both the investigators and the witnesses themselves rated the witnesses' confidence in their identification decisions 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 $r = -.16$, $p = .166$), 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 majority of investigators in Study 2 chose Jane Gibbs as their suspect, although as in Study 1 there was some variability in investigators' choice of suspect. The study was designed so that Jane Gibbs was the most likely culprit so that investigators would feel somewhat committed to their suspect choice in an attempt to mimic real life conditions. However, as real witnesses were used whose descriptions varied, suspect choices also showed some variability depending on how the culprit was described.

Investigators who chose Jane Gibbs as their suspect provided higher pre-identification probabilities of her guilt than investigators who chose someone else as their suspect. These findings replicate Study 1. As the information was designed to favor Jane Gibbs, it makes sense that those who chose her were more confident. Interestingly there were no differences in investigators' pre-identification probabilities of guilt based on whether the witness had been exposed to good or poor viewing conditions. There was an expectation that the information provided by witnesses in the poor viewing condition would be less detailed than that provided by witnesses in the good viewing condition and potentially lead investigators to be less certain in their choices; however this appeared not to have been the case. 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.

Investigators indicated that physical description contributed most to their choice of suspect. This replicates Study 1 and makes sense given that the physical description is quite salient to investigators as they personally obtain this information from the witness prior to viewing the database, whereas the other information is simply contained within the database. Fingerprint evidence was rated as second in importance, followed by prior record, alibi, and other evidence in that order.

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 estimates of the probability that their suspect was the culprit 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 do not discriminate between accurate and inaccurate witnesses (e.g., Lindsay et al., 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 did not pick up on this difference.

Rationale for Study 3

In Study 2, although witnesses were much more likely to be accurate if they had good viewing conditions versus poor viewing conditions, viewing conditions made no difference in the impact that witnesses had on investigators (although investigators were able to distinguish somewhat between witnesses who were accurate and witnesses who were not when witnesses had a good view). Two possibilities arise from the results from Study 2. First, there is the possibility that student-investigators have a tendency to weight eyewitness evidence heavily and are insensitive to viewing conditions as a factor that can compromise accuracy. Second, it is possible that investigators simply didn't pick up on the fact that some witnesses had a better view than others, as there may not have been observable differences between the two groups in spite of witnesses reporting differences. One way to get around this second possibility is to examine a factor that is readily apparent to investigators. The race of the criminal is just such a factor.

Research indicates that people are better at identifying people of their own race as opposed to those of other races, a phenomenon dubbed the own-race bias (or effect). Meissner and Brigham (2001) conducted a meta-analysis indicating that these results are robust. They found in previous studies people were 1.40 times more likely to identify someone correctly if they were same-race and 1.56 times more likely to falsely identify someone if they were cross-race. In addition, in a recent survey of potential jurors, Read and Desmarais (2006) found that the majority were aware that cross-race identifications are less reliable than same-race identifications. That is, one does not need to be an "eyewitness expert" to have knowledge of the own-race effect.

However, knowledge and behaviour are two different things. Just because people tend to be aware of cross-race identification issues does not mean that this knowledge will necessarily be

reflected in their behaviour. However, unlike in Study 2 where it is possible that the viewing conditions manipulation was too subtle to create noticeable differences in witnesses' behaviour to allow investigators to be differentially affected by witnesses who had a good versus a poor view, in Study 3, investigators were able to easily observe whether witnesses were making a cross-race identification or a same-race identification decision.

In Study 3, Caucasian witnesses viewed one of multiple Caucasian or Asian targets and then attempted an identification decision from a TA or TP lineup. It was predicted that Caucasians would demonstrate an own-race bias in that they would be more accurate when making an identification from a Caucasian lineup than from an Asian lineup as several studies have shown that Caucasian witnesses do demonstrate an own-race bias with Asian targets (e.g., Luce, 1974; Ng & Lindsay, 1994). In line with findings by Read and Desmarais (2006) it was also expected that the majority of investigators would show knowledge of the own-race bias. However, based on the results of Studies 1 and 2 and research showing that people can have knowledge of an effect but not utilize that knowledge in their decision making (e.g., Katz, 1985), it was believed that investigators would not necessarily use this information when determining how much weight to give eyewitness evidence.

Method

Participants

One hundred and seventy-six pairs of University students were recruited for this study and received either course credit or five dollars for their participation. Within each pair, one person was randomly assigned to be a witness to a videotaped crime and the other to act as an "investigator" with the constraint that the witness was always Caucasian. In addition each witness was randomly assigned to watch a video depicting one of two Caucasian or Asian

criminals and to view either a target present or a target absent 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. The video depicted a full frontal body and headshot of one of four targets (2 Caucasian and 2 Asian) walking into a room and turning toward a desk approximately 8 feet from the camera. The target spoke to a woman seated off camera who then left the room. After looking around for several seconds, the target rifled through a purse on the desk and pulled out a wallet. He removed money from the wallet and put it in his pocket. The woman then re-entered the room (although never appearing on camera) and spoke briefly to the target. The target then exited the room through the same doorway from which he entered. Each video lasts approximately 35 seconds (mean length for Caucasian targets = 34 seconds and for Asian targets = 36 seconds).

Police database. Two "police databases" were created using the same computer program as in the previous studies and included information about 13 men who could be possible suspects in the case. One database described people of Asian descent (Chinese, Japanese, and Korean) and one database described people of Caucasian descent (and one African American).

As in the previous studies, information about each suspect's general physical description, birth date, and prior criminal record was provided, along with additional information that included details regarding the suspect's vehicle, employment, and alibi for the time of the crime. Likewise, fingerprint information was included for some suspects, given as a percent probability that a fingerprint lifted from the crime scene belonged to the suspect. As in the previous studies,

the information in the database tended to favor one potential suspect: John Gibbs when the criminal was Caucasian and John Choi when the criminal was Asian. The descriptions given for each of these suspects were created based on pilot testing. The information provided for John Gibbs and John Choi is available in Appendices L and M. As the two Caucasian criminals and the two Asian criminals were similarly described, the same descriptions were used for both Caucasian and both Asian criminals.

Photo lineup. The photo lineup consisted of a set of 6 pictures taken from the front on a uniform background showing each member's head and neck, and 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. Each member of the lineup wore different clothing but no member wore the clothing that the criminal wore during the crime. 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 filled out the same two questionnaires as they did in the previous studies. The only difference was the inclusion of the following two questions at the end of the post-identification questionnaire:

1. There are situations where eyewitnesses have to identify members of their own race (e.g., a White person identifying another White person), and other situations where they must identify members of other races (e.g., an Asian person identifying a White person).

In YOUR opinion, eyewitnesses are:

a. Much more accurate when identifying members of other races than members of their own race
b. Slightly more accurate when identifying members of other races than members of their own race
c. Equally accurate in identifying members of both their own race and other races
d. Slightly less accurate when identifying members of other races than members of their own race
e. Much less accurate when identifying members of other races than members of their own race
d. I don't know

2. Please rate your confidence in your response to the previous question. (check one)
□ LOW □ MEDIUM □ HIGH

These questions were included at the end of the questionnaire so as not to affect any of the investigators' previous responses. Copies of the two questionnaires (not including these two questions) are included in Appendices B and C. Witness filled out the same questionnaire as in the previous studies. The full questionnaire is included in Appendix D.

Procedure

The procedure was identical to that of Studies 1 and 2 except that rather than being assigned to watch either a good or a poor view of the crime, witnesses were randomly assigned to view a crime committed by one of two Asian or Caucasian criminals. Investigators and witnesses were offered the same incentive as in Study 2 (i.e., the chance to be entered into a draw for $100). Five (3%) investigators failed to choose a suspect, claiming that all of the potential suspects in the police database were poor matches. For those 5 investigators, the study was stopped and both the investigator and the witness were debriefed.

Results

Pre-Identification Questionnaire

As in the previous studies, investigators filled out a questionnaire prior to conducting the lineup that asked them to indicate who they had selected as their suspect, to rate the likelihood that their suspect was guilty on a scale from 0% to 100%, to indicate whether they would arrest the suspect at this point in the investigation, and to report what factors had contributed to their selection of their suspect. Likewise, target presence was included as a factor in the majority of the pre-identification analyses to ensure that the experimenter had not affected any of these measures as she was not blind to condition.

Selection of suspect. There were no differences between the Caucasian and Asian conditions in terms of the suspect chosen. In the Caucasian target condition, 79.6% of participants chose John Gibbs (n = 82) and 20.4% chose other suspects (n = 21) whereas in the Asian target condition, 76.7% of participants chose John Choi (n = 56) and 23.3% chose other suspects (n = 17), $X^2(1) = .65$, $p = .711$, $V = .04$. The rest of the choices were scattered between the other potential suspects, although 7 Caucasian suspects (Bruce Hoffman, Tom Rabin, Callum Vinson, Hans Lindholm, Guy Fuot, Hank Ellis, and Nigel Ames) and 5 Asian suspects (Wei Chen, Ping Li, James Kim, David Huang, and Edward Wu) were not chosen by anyone.

Investigator's pre-identification probabilities suspect committed crime. Overall, investigators indicated that there was a 66.26% (SD = 17.72) chance that their suspect was guilty. A 2 (Race of Target: Caucasian, Asian) x 2 (Target: Present, Absent) x 3 (ID Decision: ID Suspect, ID Foil, No ID) ANOVA found that there were no differences in investigators' pre-identification probabilities that their suspect was guilty as a function of the target's race ($F(1, 163) = 2.25$, $MSE = 313.60$, $p = .136$, partial $\eta^2 = .01$), target presence ($F < 1$), or the witness' subsequent identification decision ($F < 1$). There were also no significant interactions (all Fs < 1.22).

The investigators' suspect choices were separated into John Gibbs versus all others when the suspect was Caucasian and John Choi versus all others when the suspect was Asian. Those who chose either John Gibbs (M = 65.21, SD = 19.65) or John Choi (M = 70.57, SD = 14.17) as their suspect provided significantly higher pre-identification probabilities that he was guilty compared to those who chose someone else as their suspect (M = 60.81, SD = 19.06, $F < 1$ and M = 63.69, SD = 14.61, $F(1, 70) = 2.90$, $MSE = 203.56$, $p = .093$, partial $\eta^2 = .04$ respectively). These results replicate Studies 1 and 2.

Decision to arrest. When the target was Caucasian, 36.89% (TP: 7 of 30, TA: 31 of 73, $z = 1.96$, $p = .050$) of investigators indicated that they would charge the suspect at this point, compared to 36.98% (TP: 13 of 40, TA: 14 of 33, $z = .86$, $p = .390$) of investigators when the target was Asian, $z = .01$, $p = .989$. 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 6.41 ($SD = 1.67$). A 2 (Race of Target: Caucasian, Asian) 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. There was a significant interaction between target presence and arrest decision, $F(1, 168) = 3.40$, $MSE = 2.66$, $p = .067$, partial $\eta^2 = .02$. When the target was absent, investigators were significantly more confident when they chose to arrest the suspect ($M = 7.04$, $SD = 1.36$) than when they chose not to arrest the suspect ($M = 5.85$, $SD = 1.77$) , $t(104) = 3.76$, $p < .001$, $d = .75$, but when the target was present, investigators were equally confident whether they chose to arrest the suspect ($M = 6.54$, $SD = 1.47$) or not ($M = 6.40$, $SD = 2.01$) , $t(68) = .28$, $p = .779$, $d = .08$. All other interactions and main effects were not significant (all other Fs < 1 except Charge Suspect: $F(1, 168) = 3.40$, $MSE = 2.66$, $p = .067$, partial $\eta^2 = .02$).

To determine whether investigators' decision to charge depended on which suspect they had chosen, John Gibbs and John Choi were separated from the other suspects chosen who were grouped into the category of "other" in order to conduct chi-square analyses. However, the decision to arrest did not depend on which suspect investigators had chosen in either the Caucasian or Asian condition, $X^2(1) = .71$, $p = .803$, $V = .04$ and $X^2(1) = .33$, $p = .394$, $V = .12$ respectively. For both Caucasian and Asian suspects, investigators were not significantly more confident when they chose John Gibbs or John Choi as their suspect than when they had selected

someone else, regardless of whether they chose to arrest the suspect ($t(36) = -1.06$, $p = .294$ and $t(25) = .62$, $p = .544$) or not ($t(63) = .99$, $p = .324$ and $t(44) = .46$, $p = .647$).

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 (Race of Target: Caucasian, Asian) 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 target race or 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, there were significant differences in rated importance between the evidence factors overall, $F(2.28, 390.56) = 129.29$, $MSE = 364.96$, $p < .001$, partial $\eta^2 = .43$. Subsequent pair-wise comparisons were carried out using the Bonferroni correction. As can be seen from Figure 7, physical description was rated as significantly more important than fingerprints ($t(174) = 6.83$, $p < .001$, $d = .92$), alibi ($t(174) = 11.75$, $p < .001$, $d = 1.47$), prior criminal record ($t(174) = 13.07$, $p < .001$, $d = 1.58$), and other evidence ($t(174) = 21.51$, $p < .001$, $d = 2.51$). Fingerprints were rated as significantly more important than alibi, prior criminal record, and other evidence ($t(174) = 4.85$, $p < .001$, $d = .54$, $t(174) = 5.76$, $p < .001$, $d = .66$, and $t(174) = 16.27$, $p < .001$, $d = 1.79$ respectively). Finally, alibi and prior criminal record were rated as significantly more important than other evidence, $t(174) = 15.05$, $p < .001$, $d = 1.63$ and $t(174) = 14.80$, $p < .001$, $d = 1.54$.

There was also an Evidence x Target Race interaction, $F(2.28, 390.56) = 129.29$, $MSE = 364.96$, $p < .001$, partial $\eta^2 = .43$. Interestingly, when the target was Caucasian, investigators reported being significantly more likely to use the physical description ($M = 43.07$, $SD = 2.11$)

than when the target was Asian (M = 35.78, SD = 2.30), $t(173) = 2.39$, $p = .018$, $d = 3.30$. There was no main effect of viewing conditions or target presence, nor were any of the other interactions significant (all Fs < 1.2).

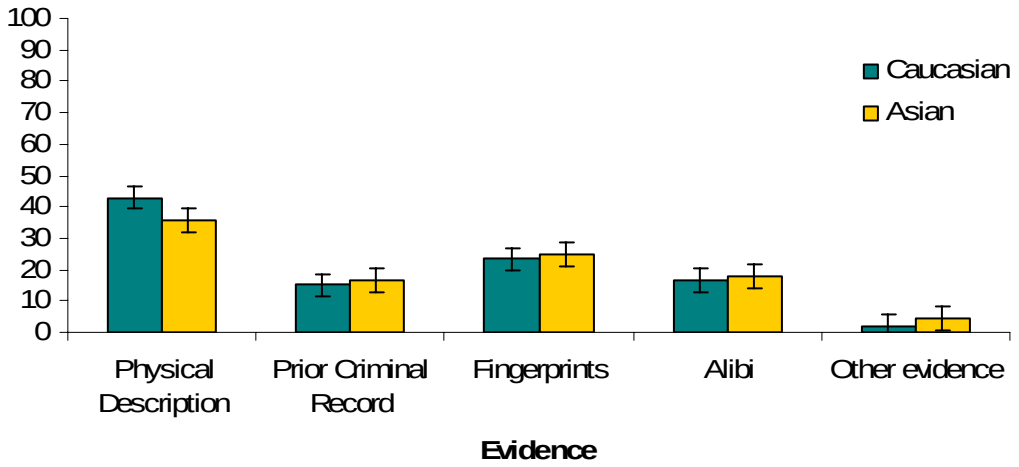

Figure 7. Mean percentage ratings of each type of evidence. The error bars represent the 95% within-subject confidence interval, appropriate for comparing the various kinds of evidence (see Masson & Loftus, 2003).

Witnesses' Identification Decisions

Identification decisions for all 176 witnesses by target race and target presence can be seen in Table 4. Although witnesses who viewed a Caucasian lineup (63.11% of the time, n = 65 of 103) were more accurate overall than witnesses who viewed an Asian lineup (52.05% of the time, n = 38 of 73), this result did not reach significance, $z = 1.45$, $p = .147$. However, looking at lineup choices separately, although witnesses were equally likely to select the guilty suspect from Caucasian and Asian lineups when the target was present (56.67% of the time, n = 17 of 30 and 57.50% of the time, n = 23 of 40 respectively, $z = .07$, $p = .947$), when the target was absent witnesses were significantly more likely to select the innocent suspect from Asian lineups

(30.30% of the time, n = 10 of 33) than Caucasian lineups (10.96% of the time, n = 8 of 73), z = 2.16, p = .031.

Likewise, there were no significant differences in witnesses' likelihood of rejecting the lineup when the target was present (Caucasian: 36.67% of the time, n = 11 of 30, Asian: 22.50% of the time, n = 9 of 40 respectively, z = 1.27, p = .204). However, when the target was absent, witnesses were significantly more likely to reject the lineup when it contained Caucasian members (65.75% of the time, n = 48 of 73) than when it contained Asian members (48.48% of the time, n = 16 of 33), z = 1.65, p = .049 (one-tailed).

Table 4. *Frequency distribution of identification choices across target absent and target present lineups in Study 3.*

	Target Absent		Target Present	
	Caucasian-Caucasian	Caucasian-Asian	Caucasian-Caucasian	Caucasian-Asian
Suspect ID	8	10	17	23
Foil ID	17	7	2	8
No ID	48	16	11	9

Witnesses were less likely to make errors when the target was Caucasian, which led us to test more participants with Caucasian targets (103) than with Asian targets (73). Even so, only 2 witnesses identified a foil when tested on a TP Caucasian lineup. As in the previous study, likely as a result of this small cell size, the data for witnesses who made foil IDs were difficult to interpret. Therefore these data have been excluded from subsequent analyses.

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 the previous studies, 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 diagnosticity of the judgment their witness had given 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 (Race of Target: Caucasian, Asian) 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 race of the target. The phase x ID decision interaction was significant, $F(1, 138) = 177.53$, $MSE = 206.21$, $p < .001$, partial $\eta^2 = .56$). Paired t-tests revealed that when the suspect was identified, investigators' guilt probabilities rose significantly from pre-lineup ($M = 69.79$, $SD = 13.71$) to post-lineup ($M = 83.76$, $SD = 17.68$), $t(57) = -6.84$, $p < .001$, $d = -.88$. If the lineup was rejected, investigators' guilt probabilities dropped significantly from pre-lineup ($M = 63.30$, $SD = 19.90$) to post-lineup ($M = 29.80$, $SD = 21.13$), $t(83) = 13.43$, $p < .001$, $d = 1.63$.

There was a main effect of phase, $F(1, 138) = 30.26$, $MSE = 206.21$, $p < .001$, partial $\eta^2 = .18$. This effect reflected the fact that, as in previous studies, the majority of witnesses (67%) made an exculpatory identification judgment resulting in an overall decline in the investigators' perceived probability that the suspect was the culprit from pre to post-ID. There was also a main

effect of identification decision, $F(1, 138) = 111.60$, $MSE = 503.56$, $p < .001$, partial $\eta^2 = .45$. Identification decision affected post-identification probabilities ($t(140) = -15.96$, $p < .001$, $d = -.88$) but not pre-identification probabilities ($F(2, 172) = 2.46$, $MSE = 308.95$, $p = .089$, partial $\eta^2 = .03$). Investigators were significantly more likely to think their suspect was the criminal when their witness identified their suspect ($M = 83.60$, $SD = 17.68$) than when the witness rejected the lineup ($M = 29.80$, $SD = 21.13$). The main effect of target race was not significant nor was the view x ID decision interaction (both Fs < 1). The overall results can be seen in Figure 8.

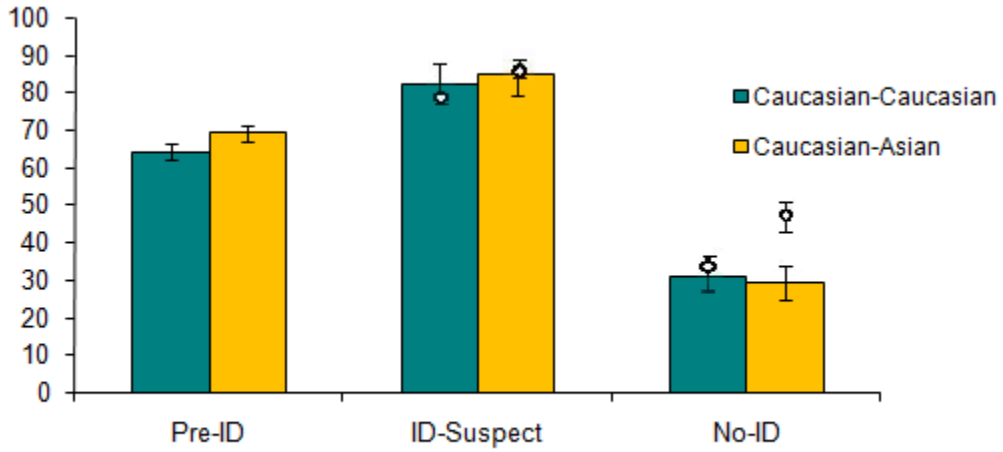

Figure 8. Investigators' percent probability suspect committed crime. The pre-lineup error bar represents a 95% within-subject confidence interval and is appropriate for comparisons between pre and post-lineup. The post-lineup error bars represent 95% between-subjects confidence intervals appropriate 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 around the individual cell means.

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, 72.00%

(TP: 13 of 17, TA: 5 of 8, $z = .66$, $p = .508$) reported they would arrest the suspect when the target was Caucasian, which was not significantly different than the 81.82% (TP: 19 of 23, TA: 8 of 10, $z = .17$, $p = .868$) who reported that they would arrest the suspect when the target was Asian, $z = .86$, $p = .391$.

In the no-ID condition, 6.78% (TP: 1 of 11, TA: 3 of 48, $z = .29$, $p = .770$) indicated that they would arrest the suspect when the target was Caucasian compared to 12.00% (TP: 0 of 9, TA: 3 of 16, $z = 1.86$, $p = .063$) who indicated that they would arrest the suspect when the target was Asian, $z = .80$, $p = .422$.

Compared to the pre-ID decisions to arrest, the frequency of investigators choosing to arrest increased significantly from 32.76% ($n = 19$ of 58) to 77.59% ($n = 45$ of 58) in the ID-suspect condition ($z = 4.69$, $p < .001$). The frequency of investigators choosing to arrest decreased significantly from 36.90% ($n = 31$ of 84) to 8.33% ($n = 7$ of 84) in the no-ID condition ($z = 5.37$, $p < .001$). These findings replicate the previous studies.

Information gain. To determine whether investigators adjusted their beliefs appropriately based on each type of identification decision, information gain analyses were conducted as outlined in the previous studies using the equations from Wells and Olson (2002) (see Appendix J). Circles on Figure 8 indicate how much investigators should have shifted their estimates based on the identification decision of the witness. Paired t-tests revealed that regardless of whether the suspect was Caucasian or an Asian, when the witness identified the suspect, investigators were extremely well calibrated, returning post-identification probabilities of guilt (Caucasian: $M = 82.52$, $SD = 20.26$, Asian: $M = 84.70$, $SD = 15.71$) that were not significantly different from the post-identification probabilities calculated based on the information gained from the

identification decision (Caucasian: *M* = 79.58, *SD* = 14.90, Asian: *M* = 85.08, *SD* = 6.71), *t*(24) = 1.14, *p* = .266, *d* = .17 and *t*(32) = -.14, *p* = .889, *d* = -.03 respectively.

However, a different pattern emerged for lineup rejections. When the lineup contained Asian members, investigators were significantly over-influenced when the lineup was rejected, lowering their probabilities of guilt (*M* = 29.04, *SD* = 21.21) significantly more than they should have based on the accuracy of the witnesses (*M* = 52.80, *SD* = 20.02), *t*(24) = -4.47, *p* < .001, *d* = -1.15. This is in contrast to lineups containing Caucasian members, where there was no evidence that investigators were over-influenced when the lineup was rejected (Investigator Post-ID rating: *M* = 30.12, *SD* = 21.28, Post-ID rating based on information gain: *M* = 32.56, *SD* = 16.50, *t*(58) = -.96, *p* = .340, *d* = -.13).

Did investigators 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. When the suspect was identified, investigators' post-identification ratings of the suspects' guilt were not significantly different following correct or incorrect identification decisions for both Caucasian (correct: *M* = 87.53, *SD* = 10.67, incorrect: *M* = 71.88, *SD* = 30.93, *t*(7.795) = -1.39, p = .202, *d* = .68) and Asian lineups (correct: *M* = 85.91, *SD* = 9.34, incorrect: *M* = 82.27, *SD* = 24.43, *t*(31) = -.62, *p* = .539, *d* = .20). Likewise, 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 both Caucasian (correct: *M* = 29.94, *SD* = 21.91, incorrect: *M* = 30.91, *SD* = 19.21) and Asian lineups (correct: *M* = 31.00, *SD* = 22.14, incorrect: *M* = 25.56, *SD* = 20.22, *t*(57) = .14, *p* = .893, *d* = .05 and *t*(23) = -.61, *p* = .549, *d* = .26 respectively). These results can be seen in Figure 9.

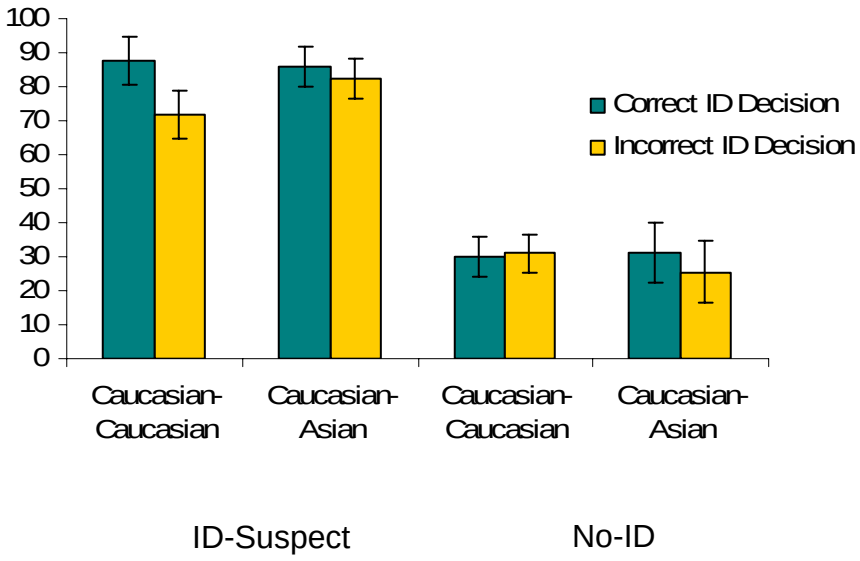

Figure 9. Investigators' percent probabilities suspect committed crime. Error bars represent 95% between-subjects confidence intervals, appropriate for comparing correct and incorrect ID decisions (see Masson & Loftus, 2003).

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 the previous studies, 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 race of target (Caucasian, Asian) was related to these ratings. However, neither of the main effects (Accuracy: $F < 1.33$, Target Race: Pillai's Trace = .21, $F(12, 75) = 1.63$, $p = .100$, partial $\eta^2 = .21$) nor the interaction ($F < 1$) were significant so these results will not be discussed further.

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. There was no relationship between eyewitnesses' self-reported confidence and their impact on investigators when the suspect was identified (Pearson's r = .17, p = .199) or the lineup was rejected (Pearson's r = -.11, p = .311).

With respect to investigators' ratings, 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 = -.30, p = .007) and a marginally significantly greater increase in investigators' ratings of guilt after the suspect was identified (Pearson's r = .25, p = .064). These findings are similar to the previous studies.

Knowledge of the Own-Race Bias

Investigators were asked to indicate whether they believed cross-race identifications on average were much more accurate, slightly more accurate, equally accurate, slightly less accurate, or much less accurate than same-race identifications. Overall, 67.50% (112 out of 166) indicated that cross-race identifications are slightly less or much less accurate than same-race identifications, 10.24% (17 out of 166) indicated that cross-race identifications are equally accurate, and 16.26% (27 out of 166) indicated that cross-race identifications are slightly more or much more accurate than same-race identifications. The remaining 6.00% (10 out of 166) responded "I don't know".

The proportion of investigators who responded correctly (slightly or much less accurate) versus incorrectly (equally, slightly, or much more accurate, I don't know) were separated into two categories in order to examine whether the proportion of investigators who responded correctly differed depending on the witnesses' identification decision or target presence. The percentage of investigators who responded correctly did not differ depending on the identification decision made by the witness (ID Suspect: 72.22% (39 out of 54), No ID: 72.22% (56 out of 79), $X^2(1) = .03$, $p = .863$, $V = .02$) or the presence of the target in the lineup (TP: 73.85% (48 out of 65) , TA: 63.37% (64 out of 101), $X^2(1) = 1.98$, $p = .159$, $V = .11$).

Investigators were also asked to rate their confidence in their answer on a scale from 1 to 3, where 1 corresponded to low, 2 corresponded to medium, and 3 corresponded to high. Investigators who accurately responded that cross-race identifications are less reliable on average relative to same-race identifications were not significantly more confident in their answers ($M = 2.36$, $SD = .59$) than investigators who responded incorrectly that investigators were equally, slightly, or much more likely to be accurate when making a cross-race identification decision ($M = 2.17$, $SD = .66$), $t(111) = -1.40$, $p = .163$, $d = .30$.

Discussion

The majority of investigators who interviewed witnesses about a Caucasian criminal chose John Gibbs as their suspect. Likewise the majority of investigators who interviewed witnesses about an Asian criminal chose John Choi as their suspect. There were no differences in the proportion of investigators who chose either John Gibbs or John Choi versus other suspects based on race. However, investigators who chose either John Gibbs or John Choi did provide significantly higher pre-identification probabilities of the suspect's guilt than investigators who chose other suspects. These findings replicate Studies 1 and 2.

Overall, investigators indicated that physical description contributed most to their choice of suspect, followed by fingerprint evidence, prior record, alibi, and other evidence in that order. These findings replicate Studies 1 and 2. Interestingly, investigators were significantly more likely to report using the physical description provided by witnesses when they were looking for a Caucasian suspect and other evidence when they were looking for an Asian suspect. There are no published studies examining whether investigators tend to rely on different types of evidence depending on the race of the culprit. However, Ellis, Deregowski, and Shepherd (1975) found that Caucasian witnesses tend to describe features that are likely to vary among Caucasian people, such as hair and eye colour, but tend not to vary as much within other cultures. As all witnesses in the current study were Caucasian, it is possible that investigators were more likely to rely on the physical description provided by witnesses when seeking a Caucasian suspect because these descriptions provided more information to help investigators differentiate between potential suspects.

The accuracy of witnesses' identification decisions was affected by whether witnesses (all Caucasian) made a same-race (Caucasian) or cross-race (Asian) identification decision. Although there were no significant differences in accuracy overall depending on whether witnesses made an identification decision about a Caucasian or an Asian lineup, witnesses were significantly more likely to select an innocent suspect from an Asian lineup than a Caucasian lineup and were significantly more likely to reject lineups that did not contain the culprit when the lineup contained Caucasian members than when it contained Asian members.

As all witnesses were Caucasian, it could be argued that the cross-race differences found in this study were not due to an own-race bias per se but rather that the Caucasian criminals were simply easier to identify than the Asian criminals. Obviously people do differ in terms of how

easy they are to recognize in a lineup; however multiple targets were used in order to reduce this possibility. A future study will test Asian witnesses under the same conditions to ensure that these results are indeed due to a cross-race effect rather than our Caucasian targets being particularly recognizable.

Investigators' post-identification probabilities replicated Studies 1 and 2 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, the race of the target did not affect the impact of the witnesses' identification decision on investigators.

In terms of information gain, when the suspect was identified, investigators were well-calibrated for both identifications of Caucasian suspects and identifications of Asian suspects. This provides further support for the conjecture provided in the discussion for Study 2 that investigators are not universally over-influenced, but rather that they are strongly influenced by eyewitness evidence and tend to be insensitive to factors that affect the accuracy of witness identifications, causing them to be over-influenced when accuracy is low to moderate but to be appropriately influenced when it is fairly high. Witnesses showed relatively high levels of accuracy when they identified the suspect, identifying Caucasian suspects and Asian suspects with accuracy levels of 68% and 70% respectively.

In further support of this idea are findings that investigators were over-influenced by witnesses who made no identification from an Asian lineup, whereas there was no evidence for over-belief when witnesses rejected a lineup containing Caucasian members. Witnesses were significantly less likely to correctly reject the lineup when it contained Asian members than

when it contained Caucasian members. This lower accuracy rate explains the over-belief found in Study 3 when lineup rejections were made from Asian lineups. As accuracy was higher when identification decisions involved Caucasian lineups, over-influence did not occur.

Finally, although investigators did not discriminate between witnesses who made a same-race versus a cross-race identification decision, the majority of investigators did report awareness of the own-race bias. Investigators who indicated that cross-race identifications are less reliable described feeling more confident about this belief than investigators who did not indicate that cross-race identifications are less reliable, however this effect did not reach significance. Thus, investigators demonstrated knowledge that cross-race identification decisions are less reliable than same-race identification decisions although this knowledge was not reflected in their decision-making. Finally, replicating findings from Studies 1 and 2 (except for suspect-IDs under good viewing conditions), investigators did not distinguish between accurate and inaccurate witnesses in terms of their post-identification probabilities of guilt.

General Discussion

This research had two central objectives. The first objective was to examine how subject-investigators are influenced by real witnesses' identification decisions 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, subject-investigators were greatly influenced by real witnesses. When witnesses were highly accurate (as was the case when they had a good view or made a same-race identification decision) 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, or when witnesses had a poor view in Study 2 or made a cross-race identification

decision in Study 3), investigators placed undue emphasis on witnesses' identification decisions. Thus, with respect to objective two, investigators for the most part failed 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. Investigators showed virtually no sensitivity to witnesses' viewing conditions, in spite of this factor having a dramatic effect on witnesses' accuracy. They also did not distinguish between same-race versus cross-race identification decisions in spite of reporting knowledge that cross-race identifications tend to be less reliable compared to same-race identifications.

Confidence measures were taken from witnesses in terms of their confidence in their identification decisions and from investigators in terms of how confident they perceived witnesses to be in their identification decisions. However, witnesses' confidence was not significantly affected by accuracy or race. Furthermore, although witnesses who had a good view indicated that they were significantly more confident that they could identify the criminal if she was in the lineup than witnesses who had a poor view, they were not significantly more confident that they were correct once they had made their identification decision.

Witnesses' confidence ratings did not tend to correlate with the impact that their identification decisions had on investigators, however, how confident investigators perceived witnesses to be in their identification decisions generally did. This is not surprising given that eyewitness confidence consistently predicts belief in the literature (e.g., Cutler et al., 1988, 1990; Lindsay et al., 1989). What is concerning is that although eyewitness confidence was correlated with eyewitness belief, accurate eyewitnesses were not significantly more confident than

inaccurate witnesses which indicates that this increase in belief was unfounded in the current research.

The results of this research should be interpreted cautiously; however, the implications of these findings are obvious. It is already known that there are hundreds of people who have been wrongfully convicted of crimes they did not commit, and that in the majority of these cases, eyewitness evidence was the major factor leading to the conviction (Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998). In the past, research efforts have focused on the role jurors play in wrongful convictions, criticizing them for placing too much credence on eyewitness testimony (Brigham & Bothwell, 1983; Lindsay et al., 1981; Wells, Lindsay, & Ferguson, 1979). However, these findings imply that the problem begins much earlier, during the police investigation.

Confirmation bias is a well known psychological process whereby people tend to seek out information that is consistent with their beliefs and disregard inconsistent information (Nickerson, 1998). In the legal system, confirmation bias is often referred to as "tunnel vision" referring to the fact that investigators who become convinced of the guilt of a suspect tend to filter incoming information in such a way as to support their belief that the suspect is indeed guilty. Incriminating evidence is elevated in significance and deemed "relevant and probative", whereas exonerating evidence is dismissed as "irrelevant or unreliable" (Findley, 2008). In many of the cases of wrongful convictions, tunnel vision has been noted to have played a role (Williamson, 2006). It is fairly easy to envision a case where investigators become convinced after a positive ID that they've found the culprit, and then selectively seek and interpret information so as to confirm that hypothesis. Unfortunately, this bias can lead investigators to inadvertently target innocent people and once these beliefs are formed, they are highly resistant

to change, even in the light of fairly conclusive disconfirming evidence, such as a positive DNA match implicating someone else (O'Brien, 2007).

It is worth noting that although the literature tends to place a greater focus on the high costs of mistaken identifications in terms of wrongful convictions, it is important to point out that investigators in the current research were also overwhelmingly compelled by lineup rejections, choosing in most cases not to arrest their suspect if he was not identified in the lineup, even if their suspect was actually guilty. In the real world, there are costs associated with mistaken non-identifications (i.e., misses) as well, as a guilty person may be eliminated as a suspect if a witness fails to identify him/her and is then released, where he/she is free to offend again.

In the future it would be beneficial to test other races to see if these results can be replicated. Of particular interest in Canada would be to test this paradigm with First Nations people as they are the most overrepresented population in Canadian prisons, making up 3% of Canadian society but over 20% of Canadian prisons (Statistics Canada, 2004). Jackiw, Arbuthnott, Pfeifer, Marcon, and Meissner (in press) examined whether the own-race bias extended to identifications involving Caucasian and First Nations people and found a significant cross-race effect; that is, Caucasian witnesses were less accurate when making identification decisions of First Nations people compared to other Caucasians and vice versa. Interestingly, they also found that both First Nations and Caucasian witnesses were more likely to select someone from First Nations lineups, suggesting that the potential for false identifications is higher for First Nations suspects. These results highlight the importance of examining whether investigators are sensitive to potential race effects in identifications involving First Nations people.

It would also be interesting to test to see whether other factors known to affect identification accuracy affect the impact of eyewitness evidence, such as age of the witness (Anastasi & Rhodes, 2006), weapon presence (Pickel, 2007), intoxication level (Dysart, 2005) or even the type of lineup procedure used (Lindsay & Wells, 1985). Other research could look at whether training investigators about factors known to affect eyewitness accuracy increases their sensitivity to these factors when determining how much weight to give eyewitness identification decisions. Developing police training methods that not only inform investigators about factors that affect eyewitness accuracy, but help investigators apply this information appears to be particularly important considering the findings of Study 3, where investigators demonstrated knowledge of the own-race bias and yet still did not use it in their decision making.

Other ideas for future research include exploring the relationship between the witnesses' behaviour during the interview and whether or not the investigator selects the intended target (and with what confidence). As well, it would be interesting to ask investigators to provide estimates of the reliability of various kinds of evidence, and then compare those values to how influenced they are by those kinds of evidence in a particular case.

The major limitation to this research is that the investigators in the current studies, as well as in Dahl et al.'s (2006) studies were undergraduate students rather than actual investigators. The next big step is to extend this research to real police investigators. It is possible that real investigators will not be as swayed by eyewitness evidence as the student-investigators in these studies. The student-investigators in the current research had exposure to only one witness which may have limited their ability to distinguish between good and poor witnesses. Real investigators have extensive experience with witnesses and may be better able to tell when a witness is or is not credible and may 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. There is also the possibility that the "guilty bias" that investigators can show during interrogations (e.g., Meissner & Kassin, 2002) might actually reflect increased skepticism generally, and extend to their interactions with witnesses as well.

Another issue in using undergraduate students rather than real investigators relates to participants' knowledge of the legal standard that must be met (i.e., reasonable grounds) in order to arrest a suspect. Undergraduates in the current research were not informed of this legal standard. Instead, they were asked simply whether they would arrest the suspect at this point in the investigation based on whether or not they believed their suspect to be guilty. Real investigators on the other hand, would have knowledge of the legal standard of reasonable grounds to arrest and may show differences in their arrest decisions as a result.

Turtle (personal communication, winter, 2005) conducted a small-scale survey of police officers to see how they said their belief in the guilt of a suspect would change based on the various identification decisions an eyewitness might make and found that they tended to indicate that they did not put much credence in eyewitness decisions. However, behavior does not always follow from people's reports on what they say they will do (Katz, 1985). In fact, Dahl (personal communication, winter, 2005) conducted a pilot study using surveys and found that people indicated that they would not be overly swayed by eyewitness evidence, and yet when she actually conducted the study, she found that they did not actually behave that way, instead being *very* compelled by eyewitness accounts. In any case, it is impossible to know how investigators will behave unless they too are tested under this paradigm.

The Need for Theoretical Perspectives on How Investigators are affected by Witnesses' ID Decisions

A common criticism of research in the area of Psychology and Law is that although it has useful applied value, it fails to inform theory (Johnson, 1993). It is true that the main focus of the current research was to look at applied rather than theoretical issues behind investigator decision making, specifically, whether the degree to which investigators are influenced by eyewitness identification evidence calibrates well with how they should be based on eyewitness accuracy. However, research on investigator decision making could be used to test and refine current theories of decision making and help in the development of models to explain investigator decision making processes. To date, attempts to explain the decision-making processes of people within the legal system have focused on the juror (or jury) and to a much lesser extent, the judge (e.g., Blanck, 1991). Currently, there is no published literature that has attempted to model investigator decision-making, i.e., how investigators decide whether or not to arrest or charge a suspect.

The information integration model and the story model are two major models that have been used to describe juror decision making. The information integration model was originally proposed by Anderson (1959) and applied to juror decision making by Kaplan (1977). It is an algebraic model where beliefs (in this case about the guilt or innocence of a suspect) are based on a weighted average of all of the information that is made available to jurors, including evidence presented at trial, such as eyewitness testimony, a confession, or an alibi, and extra-legal factors, such as the behaviour or appearance of the defendant or lawyers during the trial. Each piece of information is represented by a scale value between 0 and 1, where a value of 0 indicates “certain innocence” and 1 indicates “certain guilt” and is then weighted by the impact

that the evidence has, so a piece of evidence given by a highly credible source will be given a larger weight than evidence given by a less credible source. As each new piece of information is presented, it is weighted and scaled. At the end of the trial, all of these weighted values are summed and then divided by the total weight to form a single value corresponding to the juror's final belief in the guilt of the accused. This value is then compared to a minimum threshold which must be reached in order to vote guilty.

The story model on the other hand, developed by Pennington and Hastie (1981) is a cognitive model of juror decision-making based on the notion that jurors are more likely to vote guilty in cases where they can easily put together a plausible story outlining the sequence of events leading up to and including the commission of a crime. When jurors find it difficult to piece a story together such as when elements are missing, or when the story contains inconsistencies or appears implausible, jurors will be less likely to convict.

Future research could examine whether either of these models can be extended to investigators to describe legal decision making more generally. For example, investigators could be presented with multiple pieces of information varying in its level of incrimination and relevance to determine how factors associated with each of these models influence investigator decision making. Conducting interviews with investigators to elicit the reasoning behind their decisions to arrest a suspect or not could also help establish whether investigators tend to follow an information integration approach, a story model approach, or whether their behaviour could best be described by a different model. Thus, although the current line of research is applied, it could be used to test the limits and generalizeability of decision making models that have been applied to jurors to see if they can be extended to other legal decision makers as well.

Conclusion

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

Anastasi, J., & Rhodes, M. (2006). Evidence for an Own-Age Bias in Face Recognition. *North American Journal of Psychology*, *8*(2), 237-252.

Anderson, N. H. (1959). Test of a model for opinion change. *Journal of Abnormal & Social Psychology, 59*, 371-381.

Blanck, P. D. (1991). What empirical research tells us: Studying judges' and juries' behavior. *American University Law Review, 40,* 775-804.

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.

Bregman, N. J., & McAllister, H. A. (1987). Perceived innocence or guilt: Role of eyewitness identification and fingerprints. *Southern Psychologist, 3*(2), 49-52.

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.

Charman, S. D., & Wells, G. L. (2007). Eyewitness lineups: Is the appearance-changes instruction a good idea? *Law and human behavior*, *31*(1), 3-22.

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.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155-159.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law & Human Behavior, 14*(2), 185-191.

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law & Human Behavior, 12*(1), 41-55.

Dane, F. C. (1985). In search of reasonable doubt: A systematic examination of selected quantification approaches. *Law and Human Behavior, 9*(2), 141-158.

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.

Dysart, J. E. (2005). Intoxicated witnesses: Exploring the effects of alcohol on identification accuracy (Doctoral dissertation, Queen's University, 2005). *Dissertation Abstracts International, 66(2-B),* 1218.

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of White and Black faces by White and Black subjects. *International Journal of Psychology, 10*, 119–123.

Findley, K. (n.d.). *The problem of tunnel vision in criminal justice.* Retrieved January 3, 2008 from http://www.innocenceproject.org/docs/TunnelVision_WEB.pdf.

Garrioch, L., & Brimacombe (nee Luus), C. A. E. (2001). Lineup administrators' expectations: Their impact on eyewitness confidence. *Law and Human Behavior, 25*(3), 299-314.

Gelfand, A. E., & Solomon, H. (1973). A study of poisson's models for jury verdicts in criminal and civil trials. *Journal of the American Statistical Association, 68,* 271-278.

Haber, R., & Haber, L. (2000). Experiencing, remembering and reporting events. *Psychology, Public Policy, and Law, 6*(4), 1057-1097.

Haber. L., & Haber, R. N. (2001, November). *A meta-analysis of research on lineup identification accuracy.* Paper presented at the Annual Convention of the Psychonomics Society, Orlando, Florida.

Hastie, R., (1993). Algebraic models of juror decision processes. In Hastie, R., (Ed.)., *Inside the juror: The Psychology of juror decision making*. Cambridge: Cambridge University Press.

Jackiw, L. B., Arbuthnott, K. D., Pfeifer, J. E., Marcon, J. L., & Meissner, C. A. (in press). Examining the cross-race effect in lineup identification using Caucasian and First Nations Samples. *Canadian Journal of Behavioral Science.*

Johnson, M. T. (1993). Memory phenomena in the law. *Applied Cognitive Psychology, 7*(7), 603-618.

Kaplan, M. F. (1977). Discussion polarization effects in a modified jury decision paradigm: Informational influences. *Social Psychology Quarterly, 40*(3), 262-271.

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.

Katz, J. (1985). The role of behavioral intentions in the prediction of behavior. *Journal of Social Psychology, 125*(2), 149-155.

Leippe, M. R. (1980). Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior, 4*(4), 261-274.

Leippe, M. R., Wells, G. L., & Ostrom, T. M. (1978). Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology, 63,* 345–351.

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., Lea, J. A., Nosworthy, G. J., & Fulford, J. A. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76*(6), 796-802.

Lindsay R. C. L., Wallbridge, H., & Drennan, D. (1987). Do the clothes make the man?: An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science, 19,* 463-478.

Lindsay, R. C. L. & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law & Human Behavior, 4,* 303-313.

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.

Luce, T. (1974). The role of experience in inter-racial recognition. *Personality and Social Psychology Bulletin, 1*(1), 39-41.

Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*(4), 482-489.

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.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law, 7*, 3-35.

Meissner, C. A., & Kassin, S. M. (2002). 'He's guilty!': Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*(5), 469-480.

Natter, H. M., & Berry, D. C. (2005). Effects of active information processing on the understanding of risk information. *Applied Cognitive Psychology, 19*(1), 123-135.

Ng, W., & Lindsay, R. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology*, *25*(2), 217-232.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175-220.

O'Brien, B. M. (2007). *Confirmation bias in criminal investigations: An examination of the factors that aggravate and counteract bias*. Unpublished doctoral dissertation, University of Michigan, Michigan.

Pennington, N., & Hastie, R. (1981). Juror decision-making models: The generalization gap. *Psychological Bulletin*, *89*(2), 246-287.

Pickel, K. (2007). *Remembering and identifying menacing perpetrators: Exposure to violence and the weapon focus effect*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, *27*(6), 975-986.

R v. Lifchus, (1997) 3 S.C.R. 320.

R. v. Storrey, (1990) 1 S.C.R. 241.

Read, J. D., & Desmarais, S. L. (2006). *Re-examining juror knowledge of eyewitness issues: When is expert testimony needed?* Paper presented at the annual meeting of the American Psychology-Law Society (APLS), St. Petersburg, Florida.

Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of postidentification feedback on

eyewitness identification and nonidentification confidence. *Journal of Applied Psychology*, *89*(2), 334-346.

Statistics Canada. (2004). *Juristat.* Retrieved September 20, 2007 from Statistics Canada Homepage: http://www.statcan.ca/english/freepub/85-002-XIE/0080585-002-XIE.pdf

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* [Booklet]. Washington, DC, United States Department of Justice, Office of Justice Programs.

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., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*(4), 440-448.

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.

Wells, G. L., Olson, E. A., & Charman, S. D. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied, 9*(1), 42-52.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*(5), 835-844.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and human behavior, 22*(6), 603-647.

Williamson, T. (2006). Towards greater professionalism: Minimizing miscarriages of justice. In T. Williamson (Ed.), *Investigative interviewing: Rights, research, regulation.* (pp. 147-166). Devon, United Kingdom: Willan Publishing.

Wolf, S., & Montgomery, D. A. (1977). Effects of inadmissible evidence and level of judicial admonishment to disregard on the judgments of mock jurors. *Journal of Applied Social Psychology, 7*(3), 205-219.

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology, 91*(2), 189-202.

Yarmey, A. D. (2001). Expert testimony: Does eyewitness memory research have probative value for the courts? *Canadian Psychology, 42*(2), 92-100.

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

Appendix A

Sample from database: Information contained in suspect database for John Gibbs in Study 1

Physical Description:	• DOB: Mar. 13, 1976 • Height: 5'10 • Build: medium • Eyes: blue • Hair: short, curly, brown • Caucasian
Criminal Record:	• 2 arrests, break and enter, 1996, 1998, • 1 conviction, break and enter, 1999 • Paroled Dec. 2001
Additional Investigation:	• Alibi for duration of crime: Unknown, suspect's most recent address was in neighborhood where the crime took place. • Current employment: Unemployed • Vehicles Registered: Registered owner of a white Volkswagen Rabbit.
Fingerprint on file:	• Click for information.

Appendix B

Investigator Pre ID Questionnaire

6. If you chose a suspect, what is your suspect's name? _____

2. If you chose a suspect, what number is your suspect in the lineup? _____

3. Keeping in mind that the thief may not have been among the group of potential suspects in the database you reviewed, what do you think is the probability that the suspect you chose committed the crime? Answer by writing a number between 1% (almost no chance that your suspect is the thief) and 100% (complete certainty that your suspect is the thief); 50% would indicate that you think there is a 50/50 chance that your suspect is the thief; the closer to 1% the less sure you are, and the closer to 100% the more sure you are that your suspect is the thief. Write whatever number best represents your estimate of the probability that the suspect you chose committed the crime. _____

4. Rate the relative importance of each category of evidence by assigning a percentage (out of 100%) to each category according to its importance in your choice of suspect. For example, if you thought that only one category was important in making your decision, you would assign that category 100% and each of the other categories 0%. On the other hand, if you thought that all of the categories were equally important, you would give them each 20% (100/5). You can divide up the 100% among the five categories of evidence in any way you wish, except that they should add up to 100%.

Physical description _____
prior criminal record _____
fingerprint evidence _____
alibi (whereabouts during crime) _____
other (please specify below) _____
other: _____

5. Based on the information you have obtained so far, if this were a real case would you recommend that your chosen suspect be charged and taken to trial for this crime? Yes No

If not, what additional information would you need to convince you to charge the suspect with the crime?
_____
_____
_____

6. How confident do you feel about your decision to charge/not charge the suspect? (circle one)

1	2	3	4	5	6	7	8	9	10
not at all confident				somewhat confident					extremely confident

Appendix C

Investigator Post ID Questionnaire

1. Which (if any) lineup member did the witness choose? _____

2. If the witness made an identification, how confident are you that the witness is correct in his or her decision (that is, correctly identified the thief)? (circle one)

1	2	3	4	5	6	7	8	9	10
not at all confident				somewhat confident					extremely confident

3. Now that you have administered the lineup, what is your estimate of the probability that the suspect you chose committed the crime? Answer by writing a number between 0% (absolutely no chance that your suspect is the thief) and 100% (complete certainty that your suspect is the thief). _____

4. Do you still perceive your chosen suspect as a possible thief of the crime? (circle one) Yes No

Given that, rate the relative importance of each category of evidence in making that decision, by assigning a percentage of importance (out of 100%). For example, if you thought that only one category was important in making that decision, you would assign that category 100% and each of the other categories 0%. On the other hand, if you thought that all of the categories were equally important, you would give them each 16.67% (100/6). You can divide up the 100% among the six categories of evidence in any way you wish, except that they should add up to 100%.

Physical description _____
prior criminal record _____
fingerprint evidence _____
alibi (whereabouts during crime) _____
eyewitness decision _____
other (please specify below) _____
other: _____

5. Based on the information you have obtained so far, if this were a real case would you recommend that your suspect be charged and taken to trial for this crime? Yes No

If not, what additional information would you need to convince you to charge the suspect with the crime?

_____
_____
_____

6. How confident do you feel about your decision to charge/not charge the suspect? (circle one)

1	2	3	4	5	6	7	8	9	10
not at all confident				somewhat confident					extremely confident

7. The next few questions ask you to estimate the probability of various responses that the witness might have made. These questions are intended to assess your beliefs regarding the probability of the witness making certain responses when the thief is present or is not present in the lineup. Please answer these questions as though you had not already administered the lineup test to the witness. That is, answer the questions as though you were making the probability estimates without knowing what response the witness actually made.

For each of the following questions, answer by writing a number between 0% (absolutely no chance) and 100% (complete certainty).

Note: For the following question, your three answers should sum to 100%

If the thief <u>is</u> in the lineup, what is the probability that the witness will:

- correctly identify the thief? _____
- mistakenly identify someone else in the lineup? _____
- mistakenly not identify anyone in the lineup? _____

Note: For the following question, your two answers should sum to 100%

If the thief is <u>not</u> in the lineup, what is the probability that the witness will:

- correctly not identify anyone in the lineup? _____
- mistakenly identify someone in the lineup? _____

Finally, some overall impressions of the witness…

8. How confident did the witness appear to you during the interview?

1	2	3	4	5	6	7	8	9	10
not at all confident				somewhat confident					extremely confident

9. How confident did the witness appear to you when making their identification decision?

1	2	3	4	5	6	7	8	9	10
not at all confident				somewhat confident					extremely confident

10. How confident did the witness appear to you overall?

1	2	3	4	5	6	7	8	9	10
not at all confident				somewhat confident					extremely confident

Appendix D

Witness Pre ID Questionnaire

1. How good of a view did you get of the criminal? very poor very good 1 2 3 4 5 6 7

2. How many seconds would you estimate that the criminal's face was in view? _____ seconds

3. How well were you able to make out the criminal's face from the video? not at all very well 1 2 3 4 5 6 7

4. What would you estimate was the closest distance between the camera-eye view and the criminal's face? 10 20 30 40 50 feet

5. How much attention were you paying to the criminal's face while viewing the video? my total attention none 1 2 3 4 5 6 7

6. How confident are you that you could identify the criminal (if he was present) from a police lineup? not confident at all very confident 1 2 3 4 5 6 7

Appendix E

Witness Post ID Questionnaire

1. At the time that you identified the person in the photospread, how certain were you that the person you identified from the photos was the criminal you saw in the video?

   not at all certain 1 2 3 4 5 6 7 totally certain

2. How easy or difficult was it for you to figure out which person in the photos was the criminal?

   extremely easy 1 2 3 4 5 6 7 extremely difficult

3. After you were first shown the photos, how long do you estimate it took you to make an identification?

   I had to look for a long time to pick him out 1 2 3 4 5 6 7 I needed almost almost no time to pick him out

4. On the basis of your memory to the criminal, how willing would you be to testify in court that the person you identified was the person in the video?

   not at all willing 1 2 3 4 5 6 7 totally willing

5. Assume that an eyewitness had about the same view of the criminal that you had from the video. Do you think that an identification by this eyewitness ought to be trusted?

   definitely should not be trusted 1 2 3 4 5 6 7 definitely should be trusted

6. To what extent do you feel that you had a good basis (enough information) to make an identification?

   no basis at all 1 2 3 4 5 6 7 a very good basis

Appendix F

Interview Instructions

The other participant in this study has just completed watching a video that depicted a crime. While you are reading these instructions the experimenter will go get the other participant (the witness) and bring them back into the room. As the police investigator, you will interview the witness to gain some information about the crime and the culprit. You will later use this information to try to find a suspect using our computer database. You can ask the witness whatever questions you would like in whatever manner you would like. Below are some suggestions for information you should ask for. However, feel free to ask any additional questions that come to mind.

To allow you to pay full attention to the witness without having to worry about writing down what they say, the experimenter will record all of the witnesses' answers for you. Thus, you do not have to write down anything that the witness says on this sheet. If you have any questions about this interview please ask the experimenter now.

<u>Suggested Information to Obtain From Witness</u>:

- description of the crime
- description of the culprit
    - e.g., height
        build
        hair color
        age
        ethnicity
        facial hair
        wearing glasses
        clothes
        identifying features (tattoos, scars)
- any weapons involved
- any violence involved
- time of day

Appendix G

Notes Page

Suggested Information to Obtain From Witness:

- description of the crime

- description of the culprit

- e.g., height
  build
  hair color
  age
  ethnicity
  facial hair
  wearing glasses
  clothes
  identifying features (tattoos, scars)

- any weapons involved
- any violence involved
- time of day
-vehicle involved

Appendix H

Instructions for Trying to Find a Suspect

Now that you have completed the interview, we would like you to try to find a suspect who you think could have committed the crime. On the computer is a list of possible suspects. Imagine that the crime committed in the video just occurred and that you are a police officer in a small town trying to find the culprit. As part of your investigation you check a database that contains information on people who have prior arrests or convictions in your area. Any of these people could be potential suspects for this crime. Of course, the person who committed the crime might not be on this list at all (for example, it might have been a criminal from outside of your jurisdiction, or a local criminal who had never previously been caught). But many crimes are committed by local repeat offenders, so there is a good chance that the culprit will be in the database. Using the information you obtained from the witness interview, explore the database and decide if it includes a person whom you believe should be investigated as the suspect in this crime.

In addition to basic information about each potential suspect's appearance and prior arrest record, the database includes other kinds of information that may be helpful to you (e.g., information about alibis and about fingerprints). Imagine that a partial fingerprint was found at the scene of the crime. This fingerprint was then entered into the computer database for analysis. The computer will analyze this partial print and compare it to the fingerprints of each of the possible suspects. Some suspects may not have fingerprint information on file, but for those who do, the computer will give you the % probability that the suspect's fingerprint matches the one taken from the scene.

Take as much time as you need to make your decision. Feel free to use all of the information available to you to decide whether the person who committed the crime is one of the suspects listed here. Paper has been provided for you to make notes as needed. Please examine all of the potential suspects before you make your decision. For some of the suspects it may be quite easy to determine whether they could have committed the crime just based on their physical appearance, prior record, or alibi (where they were when the crime was committed). Other suspects may be quite similar to the description of the person who committed the crime and you may have to examine the additional evidence carefully to decide if they could have committed the crime or not.

<u>Making your decision:</u>

After going through all of the potential suspects once, the list of all of the suspects will come up and at this point you have a second opportunity to examine each of the suspects. Once you have made your decision whether or not the person who has committed the crime is in this list of suspects, click on the link to make your decision.

If you think one of these suspects could have committed the crime you can click on your suspect's name to make your decision. Once you have clicked on the suspect's name <u>your decision is final</u>. You will not be able to change your mind and go back. When you have made your decision a photo of your suspect will appear on the computer screen. Please make a note of

what your suspect looks like. Once you have made note of your suspect's appearance, click on the link to go to the lineup. Your suspect's picture will be included in this computer-generated lineup. Please make a note to yourself of where your suspect's picture is in the lineup. Once you have examined the lineup please notify the experimenter.

If you do not think that the person who committed the crime is one of these suspects then click on the link that says "none of the above." Once you have made this decision please notify the experimenter.

Appendix I

Lineup Checklist

1. Sit witness in front of computer and tell them that you will be administering a photo lineup.

2. Tell them that you are required to tell them that: “The culprit seen in the video may not be present in this lineup. If you do not think the culprit is shown here please say “Not Present” when I ask you for your identification decision.”

3. Press the button to turn on the computer screen.

4. Ask the witness to look over the photos carefully and to let you know when they have made their identification decision.

5. Ask them for their identification decision.

6. Write their decision down here: _____

7. Ask the witness any additional questions you feel would be helpful.

8. Notify the experimenter when you are done.

Appendix J

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

Appendix K

Sample from database: Information contained in suspect database for Jane Gibbs in Study 2

Physical Description:	• DOB: Mar. 13, 1979 • Height: 5'9 • Build: slender • Eyes: brown • Hair: long, wavy, brown • Caucasian
Criminal Record:	• 2 arrests, break and enter, 1996, 1998, • 1 conviction, break and enter, 1999 • Paroled Dec. 2001
Additional Investigation:	• Alibi for duration of crime: Unknown, suspect's most recent address was in neighborhood where the crime took place. • Current employment: Unemployed • Vehicles Registered: Registered owner of a white Volkswagen Rabbit.
Fingerprint on file:	• Click for information.

Appendix L

Sample from database: Information contained in suspect database for John Gibbs in Study 3

Physical Description:	• DOB: Mar. 13, 1981 • Height: 6'1 • Build: medium • Eyes: brown • Hair: short, brown • Caucasian
Criminal Record:	• 2 arrests, break and enter, 1996, 1998, • 1 conviction, break and enter, 1999 • Paroled Dec. 2001
Additional Investigation:	• Alibi for duration of crime: Unknown, suspect's most recent address was in neighborhood where the crime took place. • Current employment: Unemployed • Vehicles Registered: Registered owner of a white Volkswagen Rabbit.
Fingerprint on file:	• Click for information.

Appendix M

Sample from database: Information contained in suspect database for John Choi in Study 3

Physical Description:	• DOB: Mar. 13, 1985 • Height: 5'9 • Build: medium • Eyes: brown • Hair: black, short • Chinese
Criminal Record:	• 2 arrests, break and enter, 1996, 1998, • 1 conviction, break and enter, 1999 • Paroled Dec. 2001
Additional Investigation:	• Alibi for duration of crime: Unknown, suspect's most recent address was in neighborhood where the crime took place. • Current employment: Unemployed • Vehicles Registered: Registered owner of a white Volkswagen Rabbit.
Fingerprint on file:	• Click for information.