

2015

Pre-feedback eyewitness statements: proposed safeguard against feedback effects on evaluations of eyewitness testimony

Laura Smalarz
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>



Part of the [Legal Studies Commons](#), and the [Psychology Commons](#)

Recommended Citation

Smalarz, Laura, "Pre-feedback eyewitness statements: proposed safeguard against feedback effects on evaluations of eyewitness testimony" (2015). *Graduate Theses and Dissertations*. 14471.
<https://lib.dr.iastate.edu/etd/14471>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

**Pre-feedback eyewitness statements:
Proposed safeguard against feedback effects on evaluations of eyewitness testimony**

by

Laura Smalarz

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Psychology

Program of Study Committee:
Gary L. Wells, Co-Major Professor
Stephanie Madon, Co-Major Professor
Jason Chan
Daniel Russell
Kristi Costabile

Iowa State University

Ames, Iowa

2015

Copyright © Laura Smalarz, 2015. All rights reserved.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iii
ABSTRACT	iv
CHAPTER 1 INTRODUCTION TO THE PROBLEM	1
Post-identification Feedback as a System Variable	7
Procedural Reform Efforts	8
Insufficiency of Legal Safeguards	10
CHAPTER 2 PROPOSED SOLUTION	16
The Protective Effects of Pre-Feedback Eyewitness Statements.....	19
The Corrective Effects of Pre-Feedback Eyewitness Statements	21
CHAPTER 3 MATERIALS AND METHODS.....	24
Goals	24
Methods	25
CHAPTER 4 RESULTS	31
Experimenter Judgments of the Pre-Feedback Statements.....	31
Belief Rates	32
Evaluators' Other Judgments of Eyewitnesses' Testimonies	37
CHAPTER 5 DISCUSSION	43
Why No Impairing Effect of Feedback on Evaluators' Abilities to Discriminate?	46
Does Collecting Pre-Feedback Statements Protect Against Feedback Effects on Testimony?	53
Why No Correction Effects from Showing Evaluators Witnesses' Pre-feedback Statements?	58
CHAPTER 6 CONCLUSION.....	62
REFERENCES	65
APPENDIX A. EXPERIMENTAL DESIGN FOR THE WITNESS PHASE.....	75
APPENDIX B. EXPERIMENTAL DESIGN FOR THE EVALUATOR PHASE ...	76
APPENDIX C. STILL SHOTS FROM THE WITNESSED EVENT	77

APPENDIX D. CULPRIT-ABSENT LINEUP.....	78
APPENDIX E . CULPRIT-PRESENT LINEUP	79
APPENDIX F. PRE-FEEDBACK STATEMENT QUESTIONS.....	80
APPENDIX G. TESTIMONY INTERVIEW SCRIPT	81
APPENDIX H. TESTIMONY JUDGMENT MEASURES	83
APPENDIX I. EVALUATORS' TESTIMONY JUDGMENT MEANS.....	84

LIST OF FIGURES

	Page
Figure 1 Experimenter Ratings of Witness Pre-Feedback Statements.....	30
Figure 2 Evaluator Belief of Eyewitnesses	32
Figure 3 Evaluators' Composite Judgments of Witnesses' Testimonies	37

ABSTRACT

Mistaken but highly confident eyewitness testimony has been used to convict innocent people in more than 220 criminal cases in the United States. Research has shown that confirming post-identification feedback (e.g., “Good job, you identified the suspect”) commonly given to eyewitnesses might be partially to blame for these wrongful convictions because it inflates eyewitnesses’ reports of their confidence and other testimony-relevant eyewitness reports (Stebly, Wells, & Douglass, 2014). Indeed, recent work has shown that confirming feedback given to eyewitnesses at the time of the identification ultimately impairs the abilities of evaluators to discern whether an eyewitness made an accurate or a mistaken identification (Smalarz & Wells, 2014). The present research sought to test a novel safeguard for protecting against and correcting for the effects of confirming feedback on evaluations of eyewitness testimony: the pre-feedback eyewitness statements safeguard. Some eyewitnesses, but not others, were asked a series of testimony-relevant questions about the witnessed event and their identification decision prior to receiving confirming feedback or no feedback. These pre-feedback eyewitness statements were videotaped and were later shown to some evaluators, but not others, as the evaluators made judgments about the accuracy of eyewitnesses’ testimonies. The videotaped pre-feedback statements safeguard did not appear to protect against or correct for the effects of feedback on evaluations of eyewitness testimony. Importantly, however, a number of unexpected findings emerged in the current work that have the potential to advance our understanding of how post-identification feedback influences eyewitnesses. Future directions in light of these findings are discussed.

CHAPTER 1

INTRODUCTION TO THE PROBLEM

"All the evidence points rather strikingly to the conclusion that there is almost nothing more convincing [to a jury] than a live human being who takes the stand, points a finger at the defendant, and says 'That's the one!'"

-U.S. Supreme Court Justice William J. Brennan, 1980

By the time the U.S. Supreme Court began to express concerns about the courtroom dangers of overly persuasive eyewitnesses, psychological scientists had already come to the conclusion that jurors were too accepting of eyewitness-identification testimony as evidence of guilt. In one of the first experiments demonstrating jurors' uncritical acceptance of eyewitness testimony, testimony from a single eyewitness boosted guilty verdicts from 18% in the circumstantial-evidence condition to 72% in the eyewitness-testimony condition (Loftus, 1974). Even more troubling was the finding that mock jurors' verdicts in this study were unaffected by a discrediting manipulation in which the defense proved that the eyewitness had very poor vision and that he had not been wearing his glasses at the time of witnessing: a full 68% of the mock jurors presented with this evidence still voted to convict. This revealing study and others following it led psychologists to conclude that triers-of-fact such as judges and jurors are too easily persuaded by the testimony of eyewitnesses.

Much of the blame for the over-belief problem was initially ascribed to deficiencies of the triers-of-fact. Eyewitness researchers argued that judges and jurors lack a good understanding of how memory works and that this hampers their abilities to accurately evaluate eyewitness-identification testimony (e.g., Deffenbacher & Loftus, 1982). For example, a number of studies

showed that variables that are known to affect eyewitness accuracy (e.g., retention interval, weapon presence) are not properly accounted for by people evaluating eyewitness evidence (e.g., Cutler, Penrod, & Stuve, 1988). Similarly, research on people's intuitions about eyewitness evidence suggested that jurors' common-sense notions about eyewitness memory were inconsistent with what the science showed (Brigham & Bothwell, 1983; Deffenbacher & Loftus, 1982; Yarmey & Jones, 1983). And eyewitness experts tended to agree that jurors would be better-equipped to evaluate eyewitness testimony if they were aided by the testimony of an expert (Kassin, Ellsworth, & Smith, 1989).

At the same time, some of the blame for the over-belief problem was placed on eyewitnesses, who were criticized for being too confident in their abilities to make accurate identifications (e.g., Lindsay, Wells, & Rumpel, 1981). The early research on eyewitness confidence exposed what researchers described as a "tenuous" relation between confidence and accuracy (e.g., Clifford & Scott, 1978; Deffenbacher, Brown, & Sturgill, 1978; Leippe, Wells, & Ostrom, 1978). It seemed that the confidence expressed by an eyewitness provided little information regarding the likely accuracy of the witness' identification. This apparently dismal confidence-accuracy relation became increasingly problematic as studies showed that the confidence expressed by an eyewitness is the primary determinant of whether evaluators will believe the witness (e.g., Cutler, Penrod, & Dexter, 1990; Fox & Walters, 1986; Wells, Ferguson, & Lindsay, 1981). These findings, which developed alongside an ever-growing body of work that exposed the general fallibility of eyewitness memory (e.g., Loftus, 1979) led psychologists to be somewhat pessimistic about the legal system's reliance on eyewitness evidence for solving crimes and prosecuting suspects. The advent of forensic DNA-testing in the 1990s gave new credence to these concerns because 75% of those convicted by juries, who have

now been exonerated based on DNA testing, were cases of mistaken eyewitness identification (Innocence Project, 2015).

In the late 1990's, however, a complementary argument began to emerge that placed less blame on triers-of-fact and on eyewitnesses for these miscarriages of justice and instead focused more on the idea that the legal system might be artificially inflating the believability of mistaken eyewitness-identification testimony. In a series of experiments, it was shown that false eyewitness confidence (being highly confident or certain and yet mistaken) can be easily created by giving witnesses confirming feedback following a mistaken identification ("Good, you identified the suspect"). In the original experiment demonstrating this post-identification feedback effect, Wells and Bradfield (1998) gave confirming feedback or no feedback to mock-eyewitnesses, all of whom had made mistaken identifications. The witnesses then answered a number of questions that are typically asked of witnesses at trial. Not only did the feedback inflate witnesses' recollections of how confident they had been at the time of the identification, but it also inflated other testimony-relevant judgments such as witnesses' recollections of how good their view was, how much attention they paid while witnessing, how easy it was for them to recognize the person, and so on. Importantly, a core set of these judgments were retrospective in nature. In other words, they required the witness to make judgments about events that occurred *before* the feedback was given. For instance, witnesses were asked about how confident they had been *at the time of the identification* and how good of a view they had *at the time of witnessing*. Hence, effects of feedback on witnesses' reports were distortions. Witnesses who received feedback tended to falsely remember having been more confident than they actually were at the time of the identification and falsely remember having had a better view than they actually did at the time of witnessing.

The size of the post-identification feedback effect on mistaken eyewitnesses' retrospective judgments is quite large. In a recent meta-analysis of 20 published studies involving more than 7,000 participant-witnesses (Stebly, Wells & Douglass, 2014), it was shown that only 6% of mistaken witnesses reported confidence levels of 80% or greater when they had not received any feedback. When witnesses had been given feedback following their mistaken identifications, however, 29% (nearly five times as many) reported 80% or greater confidence. Mistaken eyewitnesses' judgments about the quality of their view, degree of attention paid, basis for having made an identification, and so forth, are likewise strongly distorted by confirming feedback. This post-identification feedback effect has also been shown to occur among children (e.g., Hafstad, Memon, & Logie, 2004), the elderly (e.g., Neuschatz et al., 2005), for lineup rejections as well as identifications (e.g., Semmler, Brewer, & Wells, 2004) and with actual eyewitnesses to serious crimes (Wright & Skagerberg, 2007).

The discovery of the post-identification feedback effect marked a shift in psychologists' thinking about how mistaken identifications lead to wrongful convictions. Rather than attributing over-belief of eyewitnesses to deficiencies of the triers-of-fact or to eyewitness overconfidence, it was now viewed at least partially as the result of actions taken by members of legal system (i.e., law enforcement officials). This realization sparked concern among psychologists for a number of reasons. First, psychologists reasoned that the majority of real-world eyewitnesses likely receive some form of post-identification feedback before trial. When eyewitnesses make an identification in the presence of a non-blind lineup administrator (i.e., a police officer or detective who knows the identity of the suspect), for example, they will likely learn immediately following their identification whether or not they identified the suspect. A recent nationwide survey of eyewitness procedures in law enforcement agencies in the United States ($n = 1,377$)

reported that only 31% of law enforcement agencies use double-blind lineup procedures in which the person administering the lineup does not know which lineup member is the suspect and which are merely fillers (Police Executive Research Forum, 2013). Hence, in approximately 70% of police agencies nationwide, the standard practice is one in which witnesses have the opportunity to receive feedback immediately following their identification decisions. But even in jurisdictions that have adopted psychological scientists' call for double-blind lineup procedures (e.g., New Jersey, North Carolina, Ohio), witnesses will likely infer whether or not they identified the suspect based on whether the suspect is indicted for the crime. Thus, psychologists were probably correct in speculating that feedback occurs in the vast majority of real-world eyewitness cases.

Another reason the post-identification feedback effect sparked concern among eyewitness researchers was because of the presumption that post-identification feedback given to eyewitnesses would undermine the abilities of fact-finders to determine whether the eyewitness had made an accurate or mistaken identification. Research has shown that people naturally rely on witnesses' statements about their confidence, view, and the degree of attention they paid while witnessing to make judgments about whether to believe the witness (e.g., Bradfield & Wells, 2000). And, under the right conditions, these statements can be valid cues to accuracy: witnesses who make accurate identifications tend to report that they were more confident, had a better view, paid more attention, etc. than do witnesses who make mistaken identifications (e.g., Bradfield, Wells, & Olson, 2002). However, feedback has been shown to inflate the retrospective self-reports of mistaken eyewitnesses more than it inflates the retrospective self-reports of accurate eyewitnesses (e.g., Bradfield, Wells, & Olson, 2002; Steblay et al., 2014). As a result, psychologists reasoned that feedback would undermine the diagnostic value of the cues on which

fact-finders rely to evaluate accuracy, thereby thwarting their efforts to discriminate between accurate and mistaken eyewitness testimony¹.

Although two studies have now shown that confirming feedback given to mistaken eyewitnesses leads to higher rates of believing the mistaken witnesses (Douglass et al., 2010; MacLean et al., 2011), it was not until recently that research tested directly the presumption that feedback impairs evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony. Smalarz and Wells (2014) exposed participant-witnesses to a simulated crime and manipulated identification accuracy by giving the witnesses a lineup that either contained (accurate) or excluded (mistaken) the actual culprit. Witnesses were given confirming feedback or no feedback following their identifications. All of the witnesses then provided videotaped testimony about what they witnessed and who they identified to an interviewer who did not know whether the witnesses had received feedback or made an accurate identification. A new sample of participant-evaluators then watched these testimony videos and judged the accuracy and credibility each eyewitness. Among witnesses who were not given feedback, evaluators believed 70% of the accurate eyewitnesses and only 36% of the mistaken eyewitnesses, indicating significant discrimination. When witnesses were under the influence of confirming feedback, however, evaluators believed accurate and mistaken eyewitnesses at nearly equal rates (about 63%). Hence, confirming feedback eliminated evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony.

¹ Discrimination refers to evaluators' abilities to believe accurate eyewitnesses more than they believe mistaken eyewitnesses.

Post-identification Feedback as a System Variable

The finding that actions taken by members of the legal system can harm people's abilities to discriminate between accurate and mistaken eyewitness testimony goes to the heart of what is called the "system variable approach" to the eyewitness science. The system variable approach was originally developed in an attempt to distinguish between factors influencing eyewitness identification evidence over which the legal system has control (system variables) and factors influencing eyewitness identification evidence over which the legal system has no control (estimator variables; Wells, 1978). System variables include factors such as the timing of witness questioning, instructions given to witnesses, and the way in which a lineup is created and administered. Estimator variables include factors such as exposure time during witnessing, whether the witness and the perpetrator are of the same race or of different races, and the degree of stress experienced by the witness at the time of the event. Wells argued that system-variable research has a unique advantage over estimator-variable research—namely, in its potential for improving the legal system. Although estimator-variable research can help address the question of whether an eyewitness' identification might have been mistaken, system-variable research can be used to prevent mistaken identifications.

The original conceptualization of the system-variable approach dealt with variables that influence eyewitness-identification accuracy directly. For example, the instructions given to witnesses (Stebly, 1997) or whether the police use a show-up or a lineup procedure (Dysart & Lindsay, 2006) directly influence the likelihood of mistaken identification. Post-identification feedback, in contrast, occurs after the witness has already made an identification, and hence cannot affect identification accuracy per se. However, in a new look at the interplay between identification accuracy and eyewitness confidence, Wilford and Wells (2013) proposed a broader

definition of eyewitness-identification accuracy. Specifically, they argued that a mistaken witness who is highly confident is *more* inaccurate than a mistaken eyewitness who is not very confident. After all, fact-finders would quickly reject the testimony of a clearly unconfident eyewitness whereas testimony from a mistaken but highly confident eyewitness poses a great risk of wrongful conviction. This new construal of eyewitness accuracy, which hinges on belief of the eyewitness, broadens the scope of variables that can be considered system variables. Post-identification feedback is a system variable because it influences the believability of eyewitness testimony and is under the control of the legal system.

Procedural Reform Efforts

As Wells (1978) had originally speculated, system-variable research yielded a number of useful recommendations for improving the quality of eyewitness evidence. In 1998, a group of eyewitness scientists published a “white paper” (Wells et al., 1998), which set forth five procedural reform recommendations generated through system-variable research that could be used to prevent mistaken identifications and minimize the chances of wrongful conviction after a mistaken identification occurs. Two of these procedural reforms were presented as partial solutions to the post-identification feedback problem: the double-blind lineup and the certainty-statement procedure.

A double-blind lineup is one in which the person administering the lineup does not know which person is the suspect and which are merely fillers (Wells & Luus, 1990). The double-blind lineup provides partial protection from the post-identification feedback problem because it prevents the lineup administrator from giving the witness feedback immediately following the identification. After all, if the lineup administrator does not know who the suspect is, there is no

way to communicate to the witness at the time of the identification whether or not the witness identified the suspect.

The certainty-statement recommendation states that the lineup administrator must collect a certainty statement from the witness at the time of the identification. The rationale behind the certainty-statement recommendation is that certainty can be a useful postdictor of accuracy (Bothwell, Deffenbacher, & Brigham, 1987; Sporer et al., 1995), but only under conditions in which the witness' certainty has not been contaminated by external factors such as feedback. Hence, a certainty statement collected from an eyewitness by a double-blind administrator at the time of the identification can provide useful information regarding the likely accuracy of the witness' identification. In contrast, a certainty statement taken for the first time during testimony might reflect any number of other factors that are unrelated to the accuracy of the identification.

Although the double-blind lineup and certainty-statement recommendations are somewhat responsive to the post-identification feedback problem, they are far from complete solutions. First, the double-blind lineup procedure keeps the administrator from delivering feedback to the witness immediately following the identification, but it does not prevent against feedback that might occur later. For example, the witness might receive feedback from the case detective after leaving the room where the double-blind lineup was administered; feedback might be communicated through interactions with a co-witness; or feedback might be inferred based on the proceedings of the case, such as whether the suspect is released or indicted. Furthermore, post-identification feedback is not the only type of external factor that can distort witnesses' retrospective self-reports. For example, questioning witnesses repeatedly (Shaw, 1996; Shaw & McClure, 1996) and instructing witnesses to prepare for cross-examination (Wells, Ferguson, & Lindsay, 1981) have also been shown to inflate witnesses' confidence in their identifications (see

Smalarz & Wells, 2015). One might argue that collecting a certainty statement from the witness at the time of the identification mitigates the dangers of these other potential post-identification influences. However, even the certainty-statement recommendation provides only partial relief from the effects of post-identification feedback because it is limited to eyewitnesses' certainty reports (and not witnesses' reports of other testimony-relevant judgments, which are also distorted by feedback) and it does not require witnesses' statements be video recorded. These limitations of the certainty-statement safeguard will be discussed in further detail in Chapter 2 when the proposed safeguard of videotaping pre-feedback eyewitness statements is introduced.

Insufficiency of Legal Safeguards

In addition to the system-variable recommendations set forth by psychologists for improving the quality of eyewitness-identification evidence, the legal system has its own “safeguards” in place for dealing with potentially unreliable identification evidence: presence of counsel at the identification, motions to suppress eyewitness testimony, voir dire, and cross-examination of eyewitnesses. These safeguards come into play at different points throughout the criminal justice process, from the time the identification is elicited (presence-of-counsel safeguard) to when eyewitness evidence is presented and challenged in court (cross-examination safeguard). But how well does each of these safeguards work to protect against wrongful conviction based on mistaken eyewitness identification?

The presence-of-counsel safeguard gives criminal defendants the right to have a lawyer present during live, in-person identification procedures. Unfortunately, this safeguard does not offer protection to suspects during photo lineups (the U.S. Supreme Court ruled in *United States v. Ash*, 1973 that there is no right-to-counsel for photo identification procedures) nor does it offer

protection to suspects who have not yet been officially charged with the crime. In a recent survey of eyewitness-identification procedures in the United States, only 21.4% of law enforcement agencies reported using live lineup procedures, in contrast to 94.1% of law enforcement agencies that reported using photo lineup procedures. Moreover, among the agencies that reported using each of these procedures, an average of only two live lineups were conducted during the year preceding the survey versus an average of 41 photo lineups (Police Executive Research Forum, 2013). Hence, there are relatively few cases in which the presence-of-counsel safeguard would even apply. Surveys of defense attorneys corroborate the idea that the presence-of-counsel safeguard offers very limited protection to police suspects: defense attorneys report being present at only 5% of their clients' identification procedures (Stinson et al., 1996). Further undermining the value of the presence-of-counsel safeguard is a concern that attorneys might not be aware of what problems to look out for when they do attend their clients' identification procedures. Indeed, research indicates that attorneys are sensitive to some factors that increase the chances of mistaken identification (i.e., lineup bias), but not others (i.e., instruction bias, presentation format; Stinson et al., 1996).

Defendants also have the right to challenge the admissibility of identification evidence by filing a motion-to-suppress. This motion-to-suppress safeguard, however, applies only to identifications that involved the use of impermissibly suggestive procedures (of which feedback is one). There is no recourse for suppressing an identification that is potentially unreliable for non-suggestiveness related factors. But even cases in which the procedures were egregiously suggestive, motions-to-suppress are rarely successful in getting the eyewitness evidence suppressed (Wells, Greathouse, & Smalarz, 2011). One of the primary reasons why motions-to-suppress rarely succeed has to do with a fundamental flaw in the legal framework for evaluating

suppression motions. The framework for evaluating motions to suppress involves two prongs. First, the defense must prove that suggestiveness occurred (Prong 1), which is difficult because suggestiveness is often hidden. If the defense can prove that suggestiveness occurred, the court then assesses whether the identification evidence might still be reliable despite the suggestion by asking witnesses questions about their certainty, the quality of their view, and the degree of attention paid during witnessing (Prong 2). Only if the court determines that the eyewitness identification is in fact unreliable will the evidence be suppressed. But, as research has shown, post-identification feedback (and other forms of suggestion) inflates witnesses' retrospective reports about their certainty, view, and attention (Wells & Bradfield, 1998; Smalarz & Wells, 2015). Hence, the very presence of suggestion leads witnesses to have a higher standing on the so-called "reliability" measures, thereby virtually ensuring that identifications obtained using suggestive procedures will pass the "reliability" test and be admitted into evidence (Wells & Quinlivan, 2009). Hence, the legal standard for evaluating suggestively-obtained eyewitness evidence is fundamentally flawed in a way that severely undermines its value as a safeguard against wrongful conviction.

The voir dire safeguard is intended to help defense attorneys identify and eliminate prospective jurors who are unwilling or unable to scrutinize eyewitness testimony. But the extensiveness of questioning during voir dire varies considerably from state-to-state, and in most courtrooms, attorneys do not have the opportunity to question venire persons about their biases for or against eyewitnesses (Van Wallendael et al., 2006). Moreover, research on the predictive utility of people's attitudes toward eyewitnesses is mixed at best. For example, one study found that mock-jurors' attitudes toward eyewitnesses were not significantly correlated with their perceptions of defendant culpability in a case involving eyewitness identification evidence

(Narby & Cutler, 1994). And, because the defense is restricted to striking only a small number of potential jurors, the voir dire safeguard assumes that the propensity for jurors to be unwilling or unable to scrutinize eyewitnesses is applicable to only a small fraction of jurors.

The safeguard that is perhaps most relevant to the proposed research is the cross-examination safeguard, which occurs at the trial level and is intended to aid jurors in their assessments of eyewitness evidence (*Greene v. McElroy*, 1959). Cross-examination has been regarded as the “greatest legal engine for the discovery of truth” (Wigmore, 1970) and is considered to be a fundamental legal safeguard against wrongful conviction resulting from mistaken identification (Walters, 1985). Importantly, however, Wigmore was referring to the ability of cross examination to sort between truth tellers and liars. Truthful eyewitnesses are a special case of people who make “genuine” errors; there is no presumption that mistaken eyewitnesses are lying.

In order for cross-examination to help jurors make more accurate determinations about eyewitness evidence in cases involving feedback, two conditions must be satisfied. First, the defense attorney must be aware that the feedback occurred. In real cases, however, it is very difficult to prove that a witness received feedback. Witnesses must not only be aware of the feedback and remember having received it, but they must also be willing to report it. Research suggests that witnesses are unable to report accurately on whether and how they have been affected by feedback (Wells & Bradfield, 1998). Second, even if the witness remembers and reports having received feedback, the jury must be able to utilize this information appropriately in their determinations about the accuracy of the eyewitness’ identification.

Smalarz, Norris, and Wells (2014) recently tested whether cross-examination can help jurors make more accurate determinations about eyewitness evidence in cases involving post-

identification feedback. Participant-evaluators viewed videotaped testimonies of accurate or mistaken eyewitnesses who had either received or had not received confirming post-identification feedback. Half of the evaluators viewed a cross-examination of the witnesses after the witnesses provided testimony, whereas the other half of evaluators viewed the witnesses' testimonies only. Cross-examination involved first asking witnesses whether they received feedback. Witnesses who reported having received feedback (all witnesses who received feedback correctly reported so²) were then questioned about whether the feedback might have distorted their testimony. Specifically, witnesses were asked whether their reports of their certainty, view, and attention might have been influenced by the feedback (e.g., Do you think that being told you were right may have influenced your answer to the question about how certain you were at the time of the identification?). Witnesses who denied being influenced by the feedback were further challenged about their responses (e.g., "So you don't think that being told you were right may have made you recall having been more confident than you actually were at the time of the identification?). In the conditions that did not include cross-examination, the witnesses were never asked whether or not they received feedback nor were they challenged about their testimonies.

Although 84% of the witnesses who received feedback admitted that some or all of their responses during testimony might have been influenced by the feedback, evaluators' belief of the witnesses held steadfast. Specifically, even after viewing cross-examinations in which witnesses who received feedback were aggressively challenged about their testimonies, evaluators were

² Witnesses in this study who did not spontaneously report having received feedback were asked "So you weren't told anything about whether or not you picked the right person?" All witnesses who received feedback reported so at this time. However, the authors did not assume that the high rates of reporting feedback in this study reflect the true rates of reporting feedback in the real world. The experimental procedures used in this study involved asking witnesses about the feedback only minutes after they received it. Real eyewitnesses are often not asked about feedback until weeks, months, or years following the identification.

still equally likely to believe the testimonies of accurate and mistaken eyewitnesses. In other words, cross-examination of eyewitnesses who received feedback was unable to correct for the feedback's distorting effects on evaluations of the eyewitnesses' testimonies. This finding lends support to the concern that has long been expressed by eyewitness scientists: cross-examination is an insufficient safeguard against wrongful conviction based on mistaken eyewitness identification (Wells et al., 1998).

The fundamental problem with the cross-examination safeguard and many of the other legal remedies for cases involving feedback is that they come into play only after the witness has already been contaminated by the feedback. Psychological scientists agree that once a witness' memory has been contaminated, there is no good way to "undo" the contamination (Smalarz & Wells, 2013). And yet, before even beginning to tackle the question of whether contamination can be "undone", it must first be established that contamination has occurred. As was described previously, it is very difficult to know in real cases whether a witness has been exposed to post-identification feedback. Moreover, there are a number of other post-identification factors aside from feedback that inflate witnesses' confidence (repeated questioning, prepping for cross-examination, etc.). It is unlikely that one could determine in a given case whether some or all of these factors occurred, let alone accurately estimate the impact of each on the witness' testimony. Hence, there is no reason to be optimistic that the traditional legal "remedies" for dealing with flawed eyewitness evidence will help fact-finders make better determinations about eyewitness evidence after post-identification feedback has occurred. What can be done, then, to mitigate the damaging effects of post-identification influences such as feedback on evaluations of eyewitness testimony?

CHAPTER 2

PROPOSED SOLUTION

The current research tested a novel safeguard for protecting against the effects of post-identification feedback on evaluations of eyewitness testimony: video recording pre-feedback eyewitness statements at the time of the identification (“pre-feedback eyewitness statements safeguard”). I argue that this safeguard has the potential to restore evaluators’ abilities to discriminate between accurate and mistaken eyewitnesses who received feedback. A more limited version of this safeguard was proposed by eyewitness scientists in the white paper published by the American Psychology-Law Society and was adopted by the U.S. Department of Justice in the first national guide for law enforcement on eyewitness identification evidence in the United States (National Institute of Justice, Technical Working Group, 1999). The original recommendation (the certainty-statement recommendation; see Wells et al., 1998) stated that “A clear statement should be taken from the eyewitness at the time of the identification and prior to any feedback as to his or her confidence that the identified person is the actual culprit.” I will refer to the original recommendation as the “certainty statement” safeguard and to the currently-proposed recommendation as the “pre-feedback eyewitness statements” safeguard.

The certainty-statement safeguard and the pre-feedback eyewitness statements safeguard are based on the same basic premise, namely that collecting statements from witnesses at the time of the identification is the only way to guarantee that they have not been contaminated by

post-identification factors³. However, the value of the certainty-statement safeguard is limited in two important ways.

The first way in which the certainty-statement safeguard is limited is because it does not require that the witness' statement be video-recorded. Recent research has shown that feedback influences witnesses' testimonies in complex ways that are transmitted through testimony but that are not fully captured by witnesses' answers to paper-pencil measures (Smalarz & Wells, 2014). Thus, merely documenting on paper the witness' response to the certainty question has the potential to leave out diagnostic behavioral information. Take for example two witnesses who give the answer "Pretty certain" after being asked about their certainty in their identifications. Depending on how quickly each witness responds, the intonation expressed (e.g., "Pretty certain?" vs. "Pretty certain."), facial expressions, body language and other non-verbal behaviors, observers would likely draw very different conclusions about each witness' level of certainty. These potentially diagnostic non-verbal cues would be lost in a mere written documentation of the witness' statements. In contrast, videotaping witnesses' pre-feedback statements creates the best possible record for assessing witnesses' judgments.

The second limitation of the certainty-statement safeguard is that it requires documentation only of the witness' certainty statement. It does not require witnesses to provide statements about other testimony-relevant judgments such as how good of a view they had during witnessing, how closely they attended to the face of the person, and so forth. But the large body of research on the post-identification feedback effect has shown that witnesses' responses

³ As is specified in the original certainty-statement recommendation (Wells et al., 1998), it is imperative that the lineup be administered double-blind, such that the administrator does not know which lineup member is the suspect and which are merely fillers. Non-blind lineup administrators have been shown to influence not only who the witness identifies (Greathouse & Kovera, 2009) but also the certainty that the witness comes to express in the identification (Garrioch & Brimacombe, 2001). Thus, the pre-feedback eyewitness statements safeguard must be implemented in conjunction with double-blind lineup administration.

to all of these questions are inflated by post-identification feedback (see meta-analysis by Steblay et al., 2014). Moreover, witnesses' reports of their view and attention have been shown to influence fact finders' evaluations of eyewitnesses (Bradfield & Wells, 2000). Hence, recording only witnesses' certainty statements falls short as a safeguard given the pervasive effects of post-identification feedback on eyewitnesses' other retrospective reports. The pre-feedback eyewitness statements safeguard, in contrast, requires witnesses to answer a number of testimony-relevant questions about their identification and about the witnessed event. After all, memory does not get better with time (Ebbinghaus, 1885), so the sooner one can create a record of witnesses' reports of such matters, the better. And, by creating the record immediately at the time of the identification, one can be confident that the record constitutes an objective reflection of the witness' actual experience at the time of the identification, free from external contamination. In contrast, statements given by witnesses for the first time during testimony might reflect any number of other post-identification factors.

The current work proposes two theoretical mechanisms by which the pre-feedback eyewitness statements safeguard will help evaluators do a better job at discriminating between accurate and mistaken eyewitness testimony in cases that involved feedback. The *protective hypothesis* posits that collecting pre-feedback eyewitness statements will protect witnesses from the inflating effects of feedback. The *corrective hypothesis* posits that showing the pre-feedback eyewitness statements to evaluators along with the witnesses' testimonies will improve evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony. The details of these two proposed processes will be discussed in turn.

The Protective Effects of Pre-Feedback Eyewitness Statements

The dominant conceptualization of the post-identification feedback effect—the cue-accessibility conceptualization—postulates that eyewitnesses do not have access to strong internal cues about their confidence at the time of the identification, view and attention paid during witnessing, and other such judgments. Therefore, they tend to infer these judgments retrospectively at the time they are first questioned, using as a source of information their own behavior at the time of the identification as well as external cues such as feedback (Charman & Wells, 2012; Wells & Bradfield, 1998, 1999). This cue-accessibility conceptualization is similar to Bem’s self-perception theory (Bem, 1972). Bem posited that when internal cues are weak, people make inferences about their behavior and attitudes by observing their behavior and the context in which the behavior occurs. In the case of eyewitness identification, the eyewitness performs a behavior (making an identification) in a context that might involve post-identification feedback. Therefore, to the extent that witnesses lack access to strong internal cues about these retrospective judgments, witnesses might rely on feedback to infer those recollections.

The cue-accessibility conceptualization also predicts that feedback has stronger distorting effects on the self-reports of mistaken eyewitnesses than it has on the self-reports of accurate eyewitnesses. This is because accurate eyewitnesses are believed to have a relatively strong internal cue to accuracy deriving from their experience of actual recognition. In contrast, mistaken eyewitnesses’ cues to accuracy are weak. Therefore, it has been postulated that mistaken eyewitnesses rely more heavily on feedback to infer their retrospective judgments. Research showing stronger feedback-inflation effects among mistaken than among accurate eyewitnesses supports this interpretation (Bradfield, Wells, & Olson, 2002; Charman & Wells, 2012; Steblay et al., 2014).

According to the cue-accessibility conceptualization, increasing witnesses' access to cues about such matters as their confidence, attention paid, etc., should decrease their reliance on feedback to infer these recollections. This proposition has received support using a variety of methods. For example, asking witnesses a certainty question before giving them feedback or having them engage in private thought about their certainty or view before they receive feedback makes witnesses more resistant to feedback effects (e.g., Quinlivan et al., 2009; Wells & Bradfield, 1999). However, the extent to which these prophylactic manipulations provide strong and sustained relief from feedback inflation is unclear. In the Quinlivan et al study, for example, asking witnesses about their certainty before giving them feedback provided partial, but not full, protection from feedback effects. Moreover, although witnesses who were asked about their certainty before receiving feedback were initially more resistant to certainty inflation from feedback, the feedback effect rebounded after a one-week delay. In the Wells and Bradfield study, the prior thought manipulation moderated feedback inflation on some, but not all, of the witnesses' retrospective self-reports. The meta-analysis on the post-identification feedback effect supports the idea that prophylactic manipulations provide only partial protection from the inflating effects of post-identification feedback (Stebly et al., 2014).

It is possible that the prophylactic manipulations used in previous research provided only minimal protection against feedback inflation because they were directed at only one aspect of the witness' reports: *certainty* in the Quinlivan et al. (2009) study and either *certainty* or *view* in the Wells and Bradfield (1999) study. An advantage of the pre-feedback eyewitness statements safeguard is that it could provide stronger protection against feedback-inflation effects because it involves asking witnesses about a number of different testimony-relevant judgments, thereby strengthening witnesses' cues to many indicators of accuracy. Additionally, in the current test of

the pre-feedback eyewitness statements safeguard, witnesses will answer these questions orally to a person rather than privately on paper as has typically been done in past prophylactic experiments. The more public nature of this oral, in-person aspect of the pre-feedback questions might lead witnesses to think more deeply about their answers, thereby strengthening their memory traces for their pre-feedback views. Accordingly, it was hypothesized that collecting pre-feedback eyewitness statements would protect eyewitnesses against the distorting effects of feedback. Consequently, evaluators should be better able to discriminate between accurate and mistaken eyewitnesses who received feedback when the witnesses provided pre-feedback statements compared to when they did not.

The Corrective Effects of Pre-Feedback Eyewitness Statements

Under the right conditions, evaluators can discriminate between accurate and mistaken eyewitness-identification testimony. In the Smalarz & Wells (2014) study as well as in the follow-up cross-examination study (Smalarz, Norris, & Wells, 2014), evaluators were approximately twice as likely to believe the testimony of accurate witnesses as they were to believe the testimony of mistaken witnesses, but only when the witnesses had not received post-identification feedback. Under conditions of feedback, discrimination was eliminated. Thus, to the extent that witnesses have been contaminated by post-identification factors before trial, pre-feedback eyewitness statements represent an important source of information for triers-of-fact to determine which eyewitnesses are accurate and which are mistaken.

To the extent that evaluators can detect inflation in witnesses' self-reports during testimony, they may be less likely to be persuaded by the eyewitnesses' testimonies, instead attending to statements made by the witness during the pre-feedback eyewitness statements.

Research has shown that jurors are capable of detecting inconsistencies in eyewitness testimony (e.g., Berman & Cutler, 1996; Berman, Narby, & Cutler, 1995; Brewer, et al., 1999). And evaluations of eyewitnesses have been shown to be less favorable when at trial it is brought out that the witness' confidence increased between the time of the identification and the time of testimony (Bradfield & McQuiston, 2004; Jones, Williams, & Brewer, 2008). Jones et al. also found, however, that evaluators were more tolerant of confidence inflation when the witness provided a "reasonable" explanation for the inflation (e.g., "I was nervous at the time of the identification, but now I am confident. Since the identification, I have recalled other details of the event that have made me confident that I am correct."). Perhaps evaluators would be less likely to dismiss inflation if there were evidence of inflation not only in witnesses' confidence but also in other testimony reports (quality of view, attention paid, ease of identification, etc.). In other words, jurors might be receptive to this sort of explanation for the witness' inflated confidence, but jurors are unlikely to be persuaded by similar explanations for inflated reports on other measures (e.g., "I was nervous at the time of the identification but now I remember that I had a really good view, paid close attention, etc."). The present research provides evaluators with a record of all of these testimony judgments and will assess whether they improve evaluators' abilities to sort between accurate and mistaken eyewitnesses.

The present work also advances past research in two ways. First, this research used actual testimony from participant-eyewitnesses instead of fictitious trial transcripts, which were used in both of the prior studies examining evaluators' perceptions of confidence inflation (Bradfield & McQuiston, 2004; Jones, Williams, & Brewer, 2008). Second, to date no research has addressed the potential interplay between confidence inflation and identification accuracy on evaluations of eyewitness testimony. By manipulating identification accuracy, the current work tested whether

evaluators armed with witnesses' pre-feedback statements are better able to discriminate between accurate and mistaken eyewitnesses who have been exposed to post-identification feedback.

CHAPTER 3

MATERIALS AND METHODS

Goals

This research tested whether collecting pre-feedback statements from eyewitnesses can safeguard against post-identification feedback effects on evaluations of eyewitness testimony. Specifically, two hypotheses were tested.

Protective Hypothesis

Drawing on the cue-accessibility conceptualization, it was predicted that collecting pre-feedback eyewitness statements will protect witnesses against the inflating effects of post-identification feedback. Evaluators should therefore be better able to discriminate between accurate and mistaken eyewitnesses who received feedback when the witnesses provided pre-feedback statements compared to when they did not.

Corrective Hypothesis

Showing witnesses' pre-feedback statements to evaluators will help evaluators better discriminate between accurate and mistaken eyewitnesses who received feedback. Evaluators should therefore be better able to discriminate between accurate and mistaken eyewitnesses who received feedback when the evaluators view witnesses' pre-feedback statements along with the witnesses' testimonies than when they view the witnesses' testimonies only.

Methods

Overview

This research was conducted in two phases: the witnessing phase and the evaluator phase. For the witnessing phase, the manipulated variables were identification accuracy, collection of pre-feedback eyewitness statements, and administration of confirming feedback. All witnesses provided videotaped testimony about what they witnessed and who they identified. During the evaluator phase, these testimony videotapes were shown to a new sample of testimony evaluators. Some of the evaluators viewed the witnesses' pre-feedback statements in addition to their testimonies, whereas other evaluators viewed the testimonies only. All evaluators indicated whether they believed that the witness made an accurate or a mistaken identification and provided other judgments about the witnesses' testimonies.

Witnessing Phase

Design

The witnessing phase utilized a 2 (Identification accuracy: accurate vs. mistaken) x 2 (Pre-feedback statements: collected vs. not collected) x 2 (Post-identification feedback: confirming feedback vs. no feedback) between-subjects design. Graphical depictions of the experimental design and procedures for the witness phase can be found in Appendix A.

Participants

A total of 240 participants were recruited from the participant pool at Iowa State University. The number of witnesses required per condition was determined based on effects obtained by prior research assessing evaluators' belief in eyewitness testimony (Smalarz & Wells, 2014; Smalarz, Norris, & Wells, 2014). All participants were rewarded with course credit for their participation.

Procedures

Upon entering the lab, participants were given an informed-consent document and were told that the study was designed to investigate people's tendencies to form impressions of others after having viewed them for a short time. It was important that participants did not know that they would be viewing a crime because this is what typically happens with real crime eyewitnesses. After consenting to participate, participants viewed a one-and-a-half minute video of a simulated crime in which a man at an airport switches his bag with another passenger's bag. The video is clear, in color, and provides multiple views of the culprit's face (see still shots of the video in Appendix C). After viewing the video, participant-witnesses (hereafter called witnesses) were told that the bag left behind by the man contained a bomb and that the purpose of the study was actually to see if they could identify the culprit from a photo lineup. Identification accuracy was manipulated using a technique commonly employed by eyewitness researchers: including or removing the culprit from the lineup and forcing all witnesses to make identifications (e.g., Bradfield, Wells, & Olson, 2002; Charman & Wells, 2012; Smalarz & Wells, 2014; see lineups in Appendices D and E). In the culprit-absent lineup, the culprit's photo was replaced with another photo so that both lineups contained six individuals. All witnesses who made identifications from the culprit-absent lineup made mistaken identifications whereas the majority of witnesses who make identifications from the culprit-present lineup made accurate identifications.⁴

Following their identifications, half of the witnesses provided statements about their identification to the lineup administrator, who was kept blind to whether or not they had made an accurate identification (see list of questions in Appendix F). The experimenter concurrently rated

⁴ Witnesses who made mistaken identifications from the culprit-present lineup were excluded from the sample and data collection continued to reach the desired sample size.

the witness' answer to each question on a scale from 1 (not at all) to 7 (very). These ratings were later analyzed to determine whether the pre-feedback statements yielded differences between accurate and mistaken eyewitnesses. The witnesses' statements were also videotaped for use in the evaluation phase. Because of the possibility that witnesses would become committed to their pre-feedback statements and would simply provide testimony that is consistent with their prior answers (Cialdini & Trost, 1998), certain methodological features were used in an attempt to minimize commitment effects. To that end, witnesses were not informed that their pre-feedback statements were being videotaped, they provided testimony to a different experimenter than the one who administered the lineup, and the wording of the questions used during testimony were modified slightly from the wording that was used during the pre-feedback questioning. Although commitment effects are likely to occur with real eyewitnesses who provide recorded statements, the ability of the present research to test the corrective hypothesis required an attempt to minimize commitment effects.

After witnesses provided pre-feedback statements, they were randomly assigned to receive confirming feedback ("Good job! You identified the suspect.") or no feedback. The feedback was delivered orally by the experimenter. Witnesses in the no pre-feedback statements condition did not provide any statements about their identification. Instead, they were given confirming post-identification feedback ("Good job! You identified the suspect.") or no feedback immediately following their identifications.

All witnesses were directed to another room where they were met by a new experimenter who was blind to their experimental condition. This experimenter gave the witnesses an agreement form in which witnesses were asked to consent to being videotaped during a testimony interview. The experimenter then began the testimony interview, working from a

scripted series of questions that always allowed witnesses to use their own words and never cut off the witnesses' answers (see Appendix G).

At the end of the testimony interview, participants were fully debriefed according to APA ethical standards. Witnesses who provided pre-feedback eyewitness statements were informed about having been videotaped and they were given another agreement form in which they were asked to consent to the use of those videotapes in subsequent research.

Evaluator Phase

Design

The evaluator phase utilized a 3 (Pre-feedback statements: not collected vs. collected but not shown vs. collected and shown) x 2 (Identification accuracy: accurate vs. mistaken) x 2 (Post-identification feedback: no feedback vs. confirming feedback) mixed design, with identification accuracy and post-identification feedback being within-subjects variables and pre-feedback statements being a between-subjects variable. The primary dependent measure was whether or not evaluators believed that the eyewitness made an accurate identification. Graphical depictions of the experimental design for the evaluator phase can be found in Appendix B.

Participants

A total of 180 participants were recruited from the participant pool at Iowa State University. The number of evaluators required per condition was determined based on effects obtained by prior research using assessing evaluators' belief in eyewitness testimony (Smalarz & Wells, 2014; Smalarz, Norris, & Wells, 2014). Due to technical problems with the videos during some of the sessions, however, data from 18 evaluators were removed, leaving 162 evaluators for analysis. All participants were rewarded with course credit for their participation.

Procedures

After providing informed consent, each participant-evaluator viewed the videotaped testimony of four eyewitnesses (one from each accuracy-by-feedback condition, in random order). In order to deter strategic guessing regarding the accuracy of the witnesses, evaluators were led to believe that they would be viewing the testimony of five eyewitnesses and that the actual accuracy of the witnesses was unknown, therefore making it possible that they could see all mistaken witnesses, all accurate witnesses, or some combination of accurate and mistaken witnesses. Previous research using this paradigm revealed no order effects for the four tapes viewed by the evaluators (Smalarz & Wells, 2014).

Evaluators were randomly assigned to one of the three pre-feedback statements conditions. One-third of the evaluators viewed testimonies from eyewitnesses who did not give pre-feedback statements during Phase 1 (no pre-feedback statements condition; $n = 49$). The second third of evaluators viewed testimonies from witnesses who gave pre-feedback statements during Phase 1; however, the pre-feedback statements were not shown to these evaluators (statements-collected-but-not-shown condition; $n = 57$). The final third of evaluators viewed testimonies from witnesses who gave pre-feedback statements during Phase 1 and then they viewed those witnesses' pre-feedback statements (statements-collected-and-shown condition; $n = 56$). Before viewing the witness' pre-feedback statements, evaluators in this condition read the following instructions:

“It has been argued that statements taken immediately (at the time of the identification) might be better for purposes of evaluating the reliability of an eyewitness than are statements taken at a later time (such as during testimony) because of the possibility that other events might have occurred that bolstered the eyewitness's testimony. You are

about to view a video of the same witness whose testimony you just watched. This video was recorded immediately after the witness made an identification from the lineup.”

After the evaluators finished viewing each eyewitness, they indicated whether they believed the witness made an accurate or mistaken identification and they answered additional questions about their perceptions of the witness’ testimony (see measures in Appendix H).

CHAPTER 4

RESULTS

Experimenter Judgments of the Pre-Feedback Statements

In the conditions in which witnesses provided pre-feedback eyewitness statements, experimenters (blind to the identification accuracy and feedback conditions of the witness) rated witnesses' answers to each pre-feedback statement question on a scale from 1 (lower ratings) to 7 (higher ratings). A one-way MANOVA was conducted on all of the experimenters' ratings (with the *time* measure reverse-coded) to assess whether there were discernable differences between accurate and mistaken eyewitnesses during pre-feedback questioning. As predicted by the cue-accessibility conceptualization, experimenters rated the accurate witnesses' pre-feedback statements significantly higher than they rated mistaken witnesses' pre-feedback statements, $F(10, 109) = 2.14, p = .03$, partial eta-squared = .16 (Means in Figure 1), indicating that witnesses' pre-feedback statements contained diagnostic information about their accuracy.

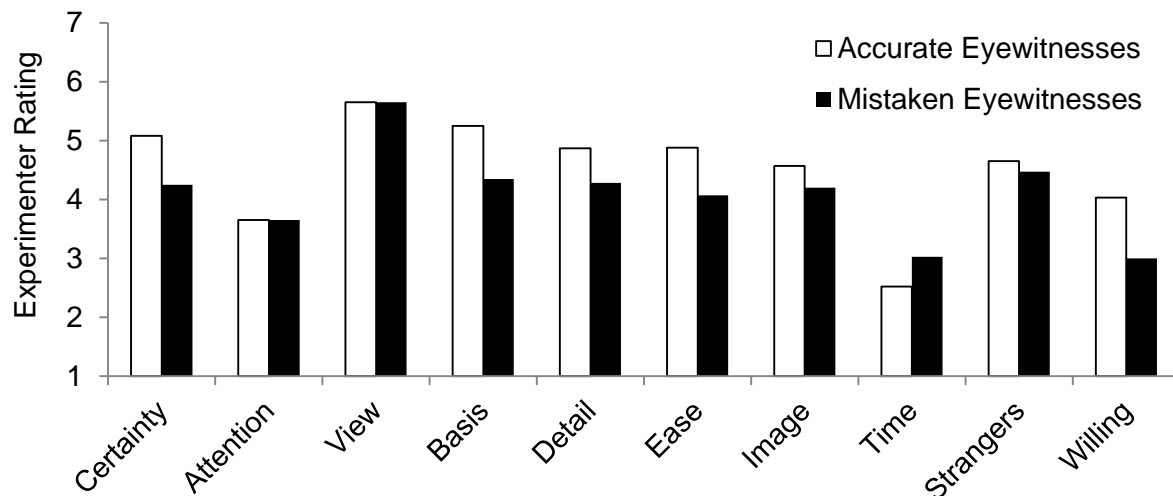


Figure 1. Experimenter Ratings of Witness Pre-Feedback. Experimenters made judgments on a scale from 1 (lower ratings) to 7 (higher ratings).

Belief Rates

The primary dependent measure was whether or not evaluators believed that the eyewitnesses had made accurate identifications. The pattern of evaluators' belief rates is shown in Figure 2. Because belief judgments were nested within evaluators (each evaluator evaluated four witnesses), observations were not independent. Therefore, a multilevel analysis was conducted using SAS PROC GLIMMIX for binary outcomes in which evaluators were identified as random effects to account for the non-independence of the dependent variable. The main effects of pre-feedback statement, accuracy, and feedback were identified as fixed effects, as were all of the higher-level interactions.

Because each evaluator evaluated four eyewitnesses, I first assessed whether evaluators' abilities to discern whether a witness made an accurate or mistaken identification (i.e., the accuracy of evaluators' judgments) improved as they viewed more witnesses. Accordingly, a statistical model was constructed in which witness order was included as a predictor of the accuracy of evaluators' judgments (along with the experimental factors and their interactions). The effect for witness order was non-significant, $F(3, 599) = .63, p = .59$. Therefore, witness order was not included in subsequent analyses.

Two statistical models were constructed to test the protective hypothesis and the corrective hypothesis, respectively. Following the tests of the two focal hypotheses, the results of a full 3 (Pre-feedback statements: not collected vs. collected but not shown vs. collected and shown) x 2 (Identification accuracy: accurate vs. mistaken) x 2 (Post-identification feedback: no feedback vs. confirming feedback) mixed-factorial analysis are reported. Effect size estimates for numeric scale measures are reported using Cohen's d , whereas effect size estimates for

differences in proportions are reported using Cohen's *h*. A *d* or an *h* of 0.2, 0.5, and 0.8 correspond to a small, medium, and large effect size, respectively (Cohen, 1988).

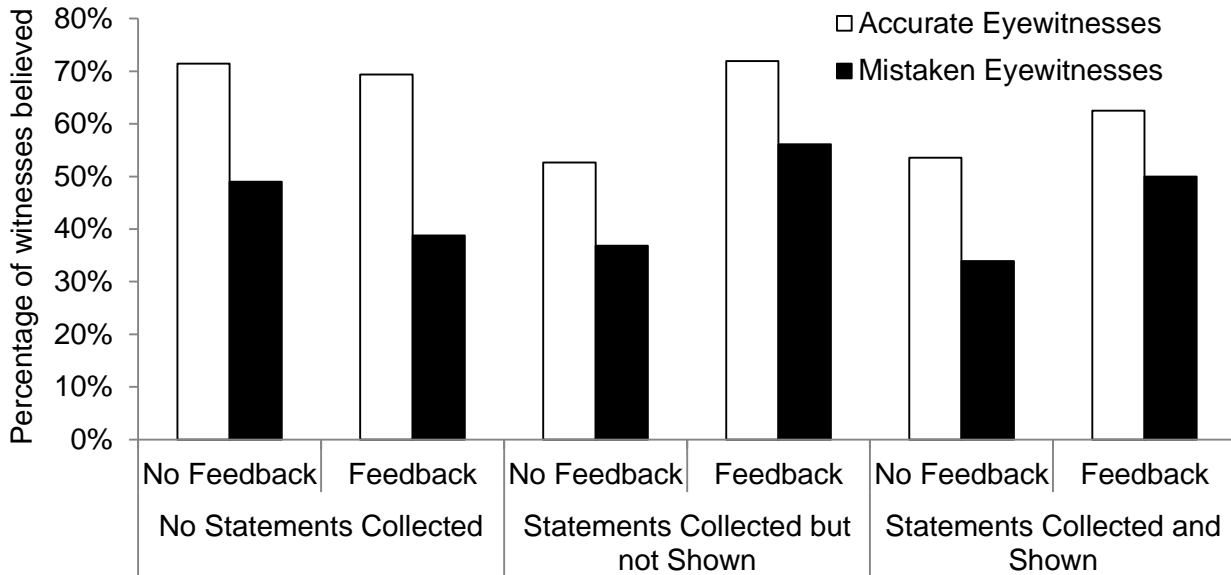


Figure 2. Evaluator Belief of Eyewitnesses

Test of the Protective Hypothesis

If collecting pre-feedback statements from eyewitnesses helps prevent their testimony from becoming distorted by feedback, then evaluators should have been better able to discriminate between accurate and mistaken eyewitnesses who received feedback when the witnesses provided pre-feedback statements compared to when they did not. Evidence of the protective hypothesis would take the form of a three-way interaction between identification accuracy, feedback, and pre-feedback statements when comparing belief rates of witnesses who gave no pre-feedback statements to belief rates of witnesses who gave pre-feedback statements (when the pre-feedback statements were not shown to evaluators). Specifically, impairment from feedback in evaluators' abilities to discriminate between accurate and mistaken testimony should be reduced among witnesses who provided pre-feedback statements compared to witnesses who provided no pre-feedback statements.

Because the protective hypothesis posits that collecting pre-feedback statements helps reduce the extent to which feedback impairs evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony, a test of the protective hypothesis requires that feedback does in fact impair discriminability in the control (no pre-feedback statements) condition. Unfortunately, however, the current work failed to replicate the finding that confirming feedback impairs evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony. Indeed, evaluators were just as able to discriminate between accurate and mistaken eyewitness testimony even when witnesses received confirming feedback, as evidenced by the non-significant two-way interaction between feedback and accuracy in the no pre-feedback statements condition, $F(1, 144) = .28, p = .60$. Moreover, this work failed to find even a main effect of feedback in the no-pre-feedback-statements condition, $F(1, 144) = .74, p = .39$. There was, however, a significant main effect of identification accuracy, indicating that the failure to replicate the feedback effect on evaluators' judgments was not due to a lack of statistical power, $F(1, 144) = 13.89, p = .0003$. Hence, I was unable to test the hypothesis that collecting pre-feedback statements from eyewitnesses reduces the extent to which feedback distorts eyewitnesses' testimonies and thereby improves evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony.

Test of the Corrective Hypothesis

If showing witnesses' pre-feedback statements to evaluators improves discrimination following feedback, then evaluators should have been better able to discriminate between accurate and mistaken eyewitnesses who received feedback when the evaluators saw witnesses' pre-feedback statements along with their testimonies compared to when evaluators saw witnesses' testimonies only. Evidence of the corrective hypothesis would take the form of a

three-way interaction between identification accuracy, feedback, and pre-feedback statements when comparing the pre-feedback statements collected-but-not-shown condition to the pre-feedback statements collected-and-shown condition. Specifically, impairment from feedback in evaluators' abilities to discriminate between accurate and mistaken eyewitnesses should be reduced among evaluators who were shown witnesses' pre-feedback statements (along with their testimonies) compared to evaluators who were shown witnesses' testimonies only.

Results did not support the corrective hypothesis. As can be discerned from Figure 2, there was not a significant three-way interaction between identification accuracy, feedback, and pre-feedback statements when comparing the pre-feedback statements collected-but-not-shown condition to the pre-feedback statements collected-and-shown condition, $F(1, 333) = .20, p = .65$. Instead, there was a significant main effect of identification accuracy across both pre-feedback statements conditions such that evaluators were more likely to believe accurate eyewitnesses (60.5%) than they were to believe mistaken eyewitnesses (44.0%), $F(1, 333) = 11.73, p = .0007$, Odds Ratio (OR) = 1.95, 95% CI (OR) = [1.33, 2.85], $h = .34$. There was also a significant main effect of confirming feedback across both pre-feedback statements conditions such that evaluators were more likely to believe eyewitnesses who received feedback (60.5%) than they were to believe eyewitnesses who received no feedback (44.0%), $F(1, 333) = 11.70, p = .0007$, Odds Ratio (OR) = 1.94, 95% CI (OR) = [1.33, 2.85], $h = .34$. There was not a significant identification accuracy x feedback interaction, $F(1, 333) = .10, p = .75$, indicating that evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony was not significantly impaired by having given confirming feedback to the eyewitnesses in the pre-feedback statement conditions.

Full-factorial Analysis of Evaluators' Belief Rates

Having tested the two focal hypotheses, a full-factorial analysis was conducted on the 3 (Pre-feedback statements: not collected vs. collected but not shown vs. collected and shown) x 2 (Identification accuracy: accurate vs. mistaken) x 2 (Post-identification feedback: no feedback vs. confirming feedback) mixed-model design.

Main Effects

Analyses indicated a significant main effect of identification accuracy such that evaluators were more likely to believe accurate eyewitnesses (64.0%) than they were to believe mistaken eyewitnesses (43.9%), $F(1, 477) = 24.82, p < .0001$, Odds Ratio (OR) = 2.27, 95% CI (OR) = [1.64, 3.13], $h = .40$. There was also a significant main effect of feedback such that evaluators were significantly more likely to believe witnesses who had received confirming feedback (58.5%) than they were to believe witnesses who had not received feedback (49.7%), $F(1, 477) = 4.76, p = .03$, Odds Ratio (OR) = 1.43, 95% CI (OR) = [1.04, 1.98], $h = .18$. There was not a significant main effect of pre-feedback statement condition, $F(2, 477) = 1.10, p = .33$.

Two-way Interactions

Only one of the three two-way interactions was significant. Specifically, there was a significant pre-feedback statements x feedback interaction, indicating that the influence of confirming feedback on evaluators' belief of eyewitnesses differed significantly by pre-feedback statements condition, $F(2, 477) = 3.59, p = .03$. Because the pre-feedback statements condition was comprised of three levels, follow-up analyses were conducted to determine which of the pre-feedback statement conditions differed in terms of the effects of the feedback on evaluators' belief rates. Among witnesses who gave no pre-feedback statements, confirming feedback did not significantly inflate evaluators' belief of eyewitnesses (60.8% no feedback vs. 54.5%

feedback), $F(1, 477) = .73, p = .39$. In contrast, feedback significantly inflated the believability of witnesses who gave pre-feedback statements when evaluators were not shown the pre-feedback statements (44.6% no feedback vs. 64.5% feedback), $F(1, 477) = 8.68, p = .003$, Odds Ratio (OR) = .44, 95% CI (OR) = [.26, .76], $h = .40$. When evaluators were shown the witnesses' pre-feedback statements, confirming feedback inflated belief of eyewitnesses, an effect that was marginally significant (43.5% no feedback vs. 56.4% feedback), $F(1, 477) = 3.59, p = .06$, Odds Ratio (OR) = .59, 95% CI (OR) = [.35, 1.02], $h = .22$. The difference in the effects of feedback on evaluators' belief of eyewitnesses was significant for the comparison between the no-statements condition and the statements-not-shown condition, $F(1, 312) = 6.89, p = .009$ and marginally significant for the comparison between the no-statements condition and the statements-shown condition, $F(1, 309) = 3.62, p = .06$. The effects of feedback on evaluators' belief of eyewitnesses did not differ significantly between the two conditions in which witnesses provided pre-feedback statements, $F(1, 333) = .57, p = .45$.

Three-way Interaction

The three-way interaction between feedback, identification accuracy, and pre-feedback statements was not significant, $F(2, 477) = .29, p = .75$.

Evaluators' Other Judgments of Eyewitnesses' Testimonies

Evaluators' other judgments of the eyewitnesses' testimonies were analyzed using a mixed-effects model on a composite measure of all of the testimony-relevant judgments (Means in Appendix I). The composite measure, which constituted a general measure of testimony credibility, was created by averaging across all of the testimony judgments, with the *additional evidence* measure being reverse-coded to create a measure of "evidence sufficiency". Hence,

higher values on the composite measure correspond to higher assessments of testimony credibility. The pattern of evaluators' judgments can be seen in Figure 3. Evaluators were again identified as random effects and accuracy, feedback, pre-feedback statements and all of higher-level interactions were identified as fixed effects.

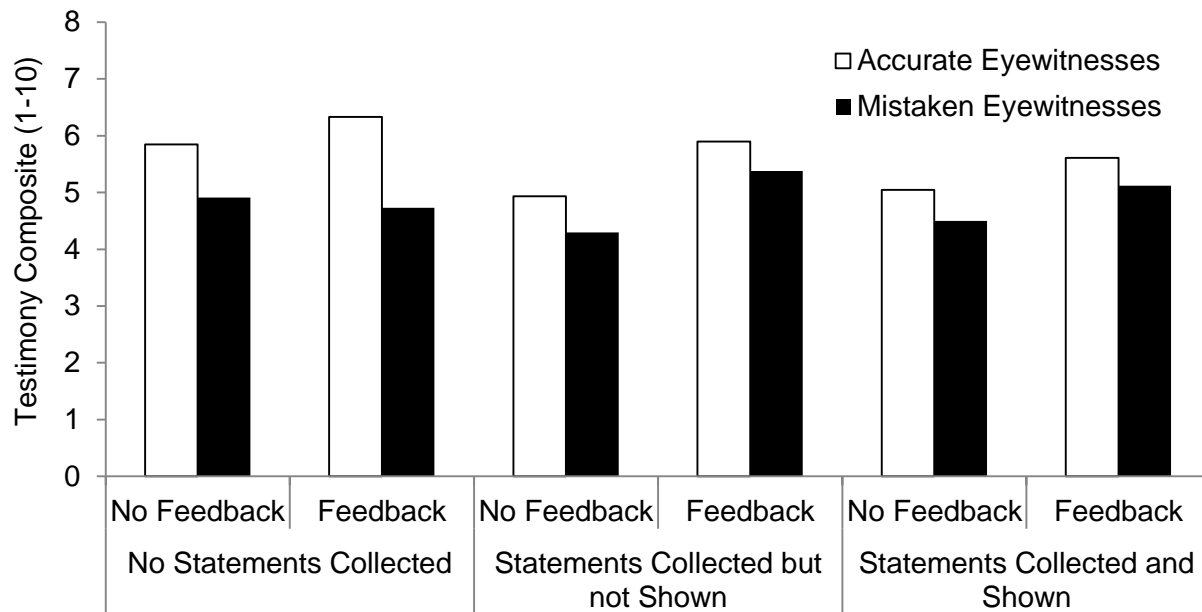


Figure 3. Evaluators' Composite Judgments of Witnesses' Testimonies. Testimony composite measure created by averaging evaluators' judgments of witnesses across the variables listed in Appendix H. Higher values indicate more favorable judgments.

Tests of Protective and Corrective Hypotheses

Recall that evidence of the protective hypothesis would take the form of a three-way interaction between identification accuracy, feedback, and pre-feedback statements when comparing the no pre-feedback statements condition to the pre-feedback statements collected-but-not-shown condition. This three-way interaction was not significant, $F(1, 404.9) = 1.32, p = .25$.

Recall also that evidence of the corrective hypothesis would take the form of a three-way interaction between identification accuracy, feedback, and pre-feedback statements when

comparing the pre-feedback statements collected-but-not-shown condition to the pre-feedback statements collected-and-shown condition. This three-way interaction was not significant, $F(1, 427.4) = .10, p = .92$. Accordingly, subsequent analyses were conducted on the full 3 (Pre-feedback statements: not collected vs. collected but not shown vs. collected and shown) x 2 (Identification accuracy: accurate vs. mistaken) x 2 (Post-identification feedback: no feedback vs. confirming feedback) mixed-factorial design.

Full Factorial Analysis of Evaluators' Judgments of Eyewitnesses

Main Effects

Analyses of evaluators' judgments of the eyewitnesses produced similar results as the belief rate analysis. There was a significant main effect of identification accuracy such that evaluators judged accurate eyewitnesses ($M = 5.59, SD = 1.86$) more favorably than they judged mistaken eyewitnesses ($M = 4.82, SD = 1.73$), $F(1, 617.6) = 33.3, p < .0001, d = .43$. There was also a significant main effect of feedback on evaluators' judgments such that evaluators judged eyewitnesses who had received confirming feedback ($M = 5.51, SD = 1.82$) more favorably than they judged eyewitnesses who received no feedback ($M = 4.90, SD = 1.80$), $F(1, 617.7) = 18.63, p < .0001, d = .34$. Finally, unlike the belief-rates analysis, there was a marginally significant main effect of pre-feedback statement condition, $F(2, 617.4) = 2.87, p = .06$. Follow-up analyses were conducted to determine which of the three pre-feedback statements conditions differed in terms of evaluators' judgments of the eyewitnesses. Evaluators tended to judge witnesses who gave no pre-feedback statements ($M = 5.46, SD = 1.90$) more favorably than they judged witnesses who gave pre-feedback statements, regardless of whether or not the statements were shown to evaluators (statements collected-but-not-shown $M = 5.12, SD = 1.86$; statements collected-and-shown $M = 5.07, SD = 1.73$). The tendency for evaluators to judge more favorably

witnesses who gave no pre-feedback statements than they judged witnesses who gave pre-feedback statements was marginally significant for the no-statements vs. statements-collected-but-not-shown comparison, $F(1, 404.9) = 3.55, p = .06, d = .18$, and statistically significant for the no statements vs. statements collected-and-shown comparison, $F(1, 402.7) = 5.19, p = .02, d = .21$. Evaluators' judgments of witnesses who gave pre-feedback statements did not significantly differ between the statements-collected-but-not-shown and the statements-collected-and-shown conditions, $F(1, 402.8) = .12, p = .73$.

Two-way Interactions

Two of the three two-way interactions were significant. First, as observed in the belief-rates analysis, there was a significant pre-feedback statements x feedback interaction, indicating that the influence of confirming feedback on evaluators' judgments of the eyewitnesses differed significantly by pre-feedback statements condition, $F(2, 617.6) = 3.34, p = .04$. Follow-up analyses were conducted to determine which of the pre-feedback statement conditions significantly differed in terms of the effects of the feedback on evaluators' judgments of the eyewitnesses. The pattern of results mirrored the pattern found in the analysis of evaluators' belief rates. Among witnesses who gave no pre-feedback statements, confirming feedback did not significantly inflate evaluators' judgments of the eyewitnesses (no feedback $M = 5.4, SD = 1.87$; confirming feedback $M = 5.5, SD = 1.95$), $F(1, 617.6) = .38, p = .54$. In contrast, feedback significantly inflated evaluators' judgments of witnesses who gave pre-feedback statements when evaluators were not shown the pre-feedback statements (no feedback $M = 4.6, SD = 1.80$; confirming feedback $M = 5.6, SD = 1.78$), $F(1, 617.6) = 19.9, p < .001, d = .56$. When evaluators were shown the witnesses' pre-feedback statements, confirming feedback likewise inflated their judgments of eyewitnesses, (no feedback $M = 4.8, SD = 1.67$; confirming feedback $M = 5.4, SD$

= 1.75), $F(1, 617.6) = 6.52, p = .01, d = .35$. The difference in the effects of feedback on evaluators' judgments of the eyewitnesses was significant for the comparison between the no-statements condition and the statements-not-shown condition, $F(1, 404.9) = 6.48, p = .01$. The comparison between the no-statements condition and the statements-shown condition, however, was not significant, $F(1, 398.8) = 1.72, p = .19$, nor was the comparison between the two conditions in which witnesses provided pre-feedback statements, $F(1, 427.4) = 1.75, p = .19$.

Second, unlike the pattern of results in the belief rate analysis, there was a significant pre-feedback statements x identification accuracy interaction. Follow-up analyses were again conducted to determine which of the pre-feedback statement conditions differed in terms of evaluators' tendencies to judge accurate eyewitnesses more favorably than they judged mistaken eyewitnesses. In all three pre-feedback statements conditions, evaluators judged the testimonies of accurate eyewitnesses significantly more favorably than they judged the testimonies of mistaken eyewitnesses, all $F_s \geq 5.04, p_s \leq .03$. However, the magnitude of the difference in evaluators' judgments of accurate versus mistaken eyewitnesses was significantly larger for the witnesses who provided no pre-feedback statements ($M_D = 1.27, SD = 1.80$) than it was for witnesses who provided pre-feedback statements (statements collected-but-not-shown $M_D = .58, SD = 1.84$; statements collected-and-shown $M_D = .52, SD = 1.71$). The difference in evaluators' judgments of accurate versus mistaken eyewitnesses was significant for the comparison between the no-statements condition and the statements-not-shown condition, $F(1, 404.9) = 4.07, p = .04, d = .38$, as well as for the comparison between the no-statements condition and the statements-shown condition, $F(1, 398.8) = 5.00, p = .03, d = .43$. The comparison between the two pre-feedback statement conditions was not significant, $F(1, 427.4) = .03, p = .86$.

The two-way feedback x identification accuracy interaction was not significant, $F(1, 617.6) = .36, p = .55$, indicating that confirming feedback did not significantly impair evaluators' abilities to discriminate between accurate and mistaken eyewitnesses.

Three-way Interaction

The three-way interaction between feedback, identification accuracy, and pre-feedback statements was not significant, $F(2, 617.64) = .81, p = .45$.

CHAPTER 5

DISCUSSION

Five main findings can be extracted from this work. First, evaluators in this experiment were able to discriminate reliably between accurate and mistaken eyewitness identification testimony. This ability to discriminate was robust across all experimental conditions; indeed, the tendency for evaluators to believe accurate eyewitnesses more than they believed mistaken eyewitnesses was not moderated by feedback condition or by pre-feedback statement condition. When examining evaluators' judgments of the eyewitnesses' testimonies, however, evaluators' tendencies to evaluate accurate eyewitnesses more favorably than they evaluated mistaken eyewitnesses was moderated by one of the experimental manipulations: namely, by pre-feedback statements condition. Although evaluators always evaluated accurate eyewitnesses more favorably than they evaluated mistaken eyewitnesses, this tendency was more pronounced when the evaluators were judging eyewitnesses who had not provided pre-feedback statements than when they were judging eyewitnesses who had provided pre-feedback statements.

Why were evaluators more likely to judge accurate eyewitnesses more favorably than mistaken eyewitnesses when the witnesses had not provided pre-feedback statements? The answer to this question is likely related to the second noteworthy finding of this work; namely, that confirming feedback did not appear to have any influence on witnesses in the no pre-feedback statements condition. Not only did feedback fail to impair evaluators' abilities to discriminate between accurate and mistaken eyewitnesses in the no pre-feedback statements conditions, but feedback did not have even a main effect on evaluators' belief rates or judgments of these eyewitnesses. This unexpected pattern of results suggests that, for some reason, the

feedback manipulation was unsuccessful in the no pre-feedback statements condition. This finding will be explored in greater detail in the section titled ‘Why no Impairing Effect of Feedback on Evaluators’ Abilities to Discriminate?’ For current purposes, however, it is important to recognize that the absence of an effect of feedback on evaluators’ perceptions of eyewitnesses in the no pre-feedback statements condition might be why evaluators showed better discrimination in that condition. After all, if feedback did not have any effect on eyewitnesses in the no pre-feedback statements condition, then evaluators would have access to relatively strong, uncontaminated, and diagnostic cues to accuracy upon which to make judgments about the eyewitnesses.

A third important finding in this work is that confirming feedback did not impair evaluators’ abilities to discriminate between accurate and mistaken eyewitnesses in the conditions in which witnesses provided pre-feedback statements. There are two possible interpretations of this finding. The first interpretation is that confirming feedback would not have impaired evaluators’ abilities to discriminate between accurate and mistaken eyewitnesses even if the witnesses had not provided pre-feedback statements. Indeed, this is what was observed in the no pre-feedback statements condition. However, there are a number of reasons why this interpretation is an unsatisfying one. First, confirming feedback had a main effect on evaluators’ belief rates and other judgments among witnesses who provided pre-feedback statements. In the control (no pre-feedback statements) condition, confirming feedback did not even have a main effect on evaluators’ perceptions, suggesting that the feedback manipulation was completely unsuccessful in that condition. The main effect of confirming feedback in the conditions in which witnesses provided pre-feedback statements, however, indicates that the feedback manipulation was successful in those conditions.

Given that the feedback manipulation was successful in inflating evaluators' perceptions of eyewitnesses in the pre-feedback statement conditions, it stands to reason that had the witnesses not provided pre-feedback statements, evaluators' abilities to discriminate between accurate and mistaken witnesses would have been reduced for witnesses who received feedback. Indeed, there is a large body of work showing that feedback has stronger distorting effects on mistaken than on accurate eyewitnesses (e.g., Bradfield, Wells, & Olson, 2002; Steblay et al., 2014), and recent research clearly shows that confirming feedback impairs evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony (Smalarz & Wells, 2014). Hence, there are reasons to believe that collecting pre-feedback statements from eyewitnesses might indeed have helped protect them from the distorting effects of the feedback, thereby improving evaluators' abilities to discriminate between accurate and mistaken eyewitnesses in the pre-feedback statements conditions. The possibility that collecting pre-feedback statements from eyewitnesses did in fact protect them from the distorting effects of post-identification feedback will be explored in more detail in the section titled 'Does Collecting Pre-Feedback Statements Protect against Feedback Effects on Testimony?'

The fourth main finding of this work is that showing witnesses' pre-feedback statements to evaluators did not improve evaluators' abilities to discriminate between accurate and mistaken eyewitnesses who received feedback. This finding will be discussed in the section below titled 'Why no Correction Effects from Showing Evaluators Witnesses' Pre-feedback Statements?'

A final finding that merits acknowledgement in the current work is that evaluators tended to judge witnesses who gave no pre-feedback statements more favorably than they judged witnesses who gave pre-feedback statements. Although this pattern of results was observed only in evaluators' judgments of the witnesses' testimonies and not in their belief ratings, the

possibility that collecting pre-feedback statements from eyewitnesses might function to decrease the persuasiveness of witnesses' testimonies is worth contemplating. Why might collecting pre-feedback statements lead witnesses to give testimony that is less persuasive to evaluators? One possibility is that witnesses who gave pre-feedback statements felt less free to inflate or embellish their testimonies than did witnesses who had not given any statements prior to the testimony interview. Research has shown that witnesses who are told to prepare for cross-examination report having been more confident in their identification than witnesses who are not told to prepare for cross-examination (Wells, Ferguson, & Lindsay, 1981). Although witnesses in this study were not told to prepare for cross-examination, they were told that they would be interviewed on videotape by another person regarding what they witnessed and who they identified. Moreover, the tone of the testimony interview, which was intended to mimic the tone of a real criminal trial, might have motivated witnesses to bolster their testimony in an effort to justify their identification decisions. This bolstering may have been reduced, however, among witnesses who provided pre-feedback statements due to their desire to maintain consistency with their prior pre-feedback statements (e.g., Cialdini & Trost, 1998).

Why No Impairing Effect of Feedback on Evaluators' Abilities to Discriminate?

The most surprising and arguably the most perplexing finding in the current work was that feedback did not impair evaluators' abilities to discriminate between accurate and mistaken eyewitnesses in the control (no pre-feedback statements) condition. Smalarz and Wells (2014) found that confirming feedback not only impaired evaluators' abilities to discriminate between accurate and mistaken witnesses but that discrimination was totally eliminated by having given confirming feedback to the witnesses. In the current work, however, evaluators were just as able

to discriminate between accurate and mistaken eyewitnesses when the witnesses received feedback as when they did not. Why did the current work fail to replicate the feedback-impairment effect on evaluators' judgments? Two possible reasons come to mind, both of which are related to methodological differences between the original Smalarz and Wells study and the current study.

The first methodological difference between the original Smalarz and Wells (2014) study and the current study was that witnesses in the original study answered the standard post-identification feedback questions (about their certainty, view, attention, etc.) on a computer before undergoing the testimony interview. In the current study, however, witnesses in the no pre-feedback statements condition did not answer any questions prior to the testimony interview. Perhaps making retrospective judgments immediately following the administration of the feedback somehow helps to solidify or strengthen the effects of feedback on witnesses' judgments. In fact, in both of the studies showing that confirming feedback influences people's evaluations of eyewitness testimony (Douglass et al., 2010; Smalarz & Wells, 2014) witnesses completed the standard post-identification feedback questions prior to undergoing the testimony interview. The current work is the first instance in which witnesses did not first answer questions about their retrospective self-reports before entering into the testimony interview.

Why might feedback's effects on eyewitness testimony depend on first having witnesses answer questions about testimony-relevant judgments immediately following the administration of the feedback? The cue-accessibility conceptualization postulates that feedback serves as an external cue upon which witnesses can rely to make inferences about retrospective judgments when their own internal cues to such judgments are weak. Perhaps at the time that feedback is administered, the feedback is a particularly salient cue to eyewitnesses, making it rather powerful

in terms of influencing witnesses' judgments. Then, when witnesses make similar judgments during testimony, the level of distortion from the feedback reflects the initial distortion levels from when the feedback was most powerful. In contrast, feedback that is temporally remote relative to the time that retrospective judgments are made might be less salient and therefore have less influence on witnesses' judgments. If this is true, then one would expect feedback to distort witnesses' retrospective judgments to a greater extent when witnesses make the judgments immediately following the feedback than when they make the judgments at a later point in time, such as during a testimony interview. This interpretation has some weaknesses, however, because feedback has been shown to have rather strong and robust effects across many different experimental methodologies and study populations (see Steblay, Wells, & Douglass, 2014) and research has shown that delaying questioning following the administration of feedback does not influence the extent to which feedback distorts witnesses' judgments (Wells, Olson & Charman, 2003). Moreover, it has even been shown that the effects of feedback can become *stronger* rather than weaker over time because witnesses' memory of internal cues weaken more quickly than will their memory for the feedback (e.g., Neuschatz et al., 2007; Quinlivan et al., 2009). Thus, the possibility that feedback's influence on eyewitness testimony might have been constrained by its timing relative to when witnesses were questioned is not an entirely appealing explanation for why feedback failed to influence evaluators' judgments in the no pre-feedback statements condition.

A second methodological difference between the original Smalarz and Wells (2014) study and the current study was that the testimony interview in the current work was considerably longer and more elaborate than the testimony interview in the original Smalarz and Wells study. Specifically, in the original study, the testimony interview was comprised of the

standard post-identification feedback questions (about confidence, view, attention, etc.), with the only additional questions being “Are you aware that the individual you identified from the photo lineup is accused of planting a bomb in the airport?” and “Please tell me in as much detail as possible exactly what it is that you witnessed in the video,” which were asked at the outset of the interview. In the current study, the interview was intended to more closely approximate the tone of an interview that witnesses might actually experience in court (see testimony interview script in Appendix G). Therefore, witnesses in this work were asked a number of questions leading up to the standard post-identification feedback questions. For example, witnesses were asked about how many people were at the airport check-in line, whether they noticed anything in particular about the people in the line, and whether the other people in the line noticed the bag-switching. They were asked to describe in as much detail as possible exactly what they witnessed at the check-in counter and to describe the person who switched the bags, as well as the bag itself. Thus, witnesses in the current study answered a number of additional questions during the testimony-interview that the witnesses in the original Smalarz and Wells study did not answer.

Why might the longer, more elaborate testimony interview used in the current study have tempered the effects of feedback on the eyewitnesses’ testimonies? There are two possibilities, one of which operates at the level of the eyewitnesses and the other that operates at the level of the evaluators. Let us first consider the possible influence of the extended interview on the witnesses themselves. During the testimony interview, witnesses in this study were questioned about specific details of the witnessed event (e.g., How many people were at the airport check-in line? Did you notice anything in particular about the people in the line? Did any of the other people in the line notice the bag-switching?). Asking witnesses these specific, detail-oriented questions might have led the witnesses to engage more fully with their memory of the witnessed

event than if they had simply been asked global questions about the quality of their view and their memory of the culprit, as was done in the original Smalarz and Wells (2014) study. Put differently, these detail-oriented questions might have promoted a form of context reinstatement by virtue of leading witnesses to mentally reconstruct the witnessed event. The memorial benefits of context reinstatement have been demonstrated in a number of studies (e.g., Fisher & Craik, 1977; Krafka & Penrod, 1985; Geiselman et al., 1986) and derive from the principle that memory retrieval is most efficient when the context of the original event is reinstated at the time of recall (i.e., *the encoding specificity principle*; Tulving & Thomson, 1973). To the extent that the detailed and event-specific questions used in the current work led to context reinstatement during the testimony interview, then witnesses might have been better equipped to accurately estimate retrospective judgments and hence might have relied less on the feedback to make inferences. According to this possibility, the witnesses might themselves have been less strongly influenced by the feedback in terms of the extent to which the feedback distorted the witnesses' responses during the testimony interview.

The second possibility as to why the longer testimony interview might have tempered the effects of feedback on evaluators' judgments operates at the level of the evaluators. In the current study, evaluators had access to an objectively larger quantity of information upon which to base their judgments of the eyewitnesses than did evaluators in the original Smalarz and Wells (2014) study because witnesses in this study answered a number of additional questions prior to answering the questions used in the original study. To the extent that eyewitnesses' testimony interviews provide evaluators with diagnostic cues to accuracy, then it stands to reason that the more thorough and lengthy the interview is, the more information evaluators have at their disposal to evaluate the witnesses and the better they will be at discriminating between accurate

and mistaken eyewitnesses. Importantly, however, it is not a longer interview per se that might help evaluators do a better job at discriminating between accurate and mistaken eyewitnesses. After all, if confirming feedback strongly distorts witnesses' answers throughout the duration of the interview, then showing evaluators a longer as opposed to a shorter interview might actually cause evaluators' judgments to be more strongly influenced by having given confirming feedback to the eyewitnesses. Showing evaluators a longer interview should help boost their abilities to discriminate between accurate and mistaken eyewitnesses only to the extent that witnesses' answers to the additional questions are reliable indicators of their accuracy. In other words, witnesses' answers to the additional questions must not have been strongly distorted by confirming feedback.

What are the chances that the additional questions asked of witnesses in the current work were not strongly distorted by confirming feedback? As described above, the additional questions tended to ask witnesses about specific details about the witnessed event. These questions are actually quite different from the standard post-identification feedback questions, which tend to involve evaluative, meta-cognitive judgments. The standard questions are evaluative in the sense that they are couched in terms of a quality judgment (e.g., the relative "goodness" or "poorness" of the judgment). For example, to answer the question "How good was your view?" witnesses must make not only an objective judgment based on characteristics of the viewing experience, but they must also make an evaluative, meta-cognitive judgment about the relative "goodness" of that view. Standard post-identification feedback questions about how much attention witnesses paid, how well they could make out facial details of the culprit during the event, how certain they were at the time of identification, how easy they found it to make an identification, how clear of an image of the culprit they had in memory, how willing they would

be to testify about their identification, how good of a basis they had for making their identification, and how good their memory is for the faces of strangers likewise require witnesses to make evaluative, meta-cognitive judgments.

There are three exceptions to these standard post-identification feedback questions that tend to involve evaluative, meta-cognitive judgments. The standard post-identification feedback questions also typically ask witnesses about how long the culprit was in view, how far away the culprit was during the witnessed event, and how long it took to make an identification from the lineup. These three judgments are arguably more objective and less meta-cognitive in nature than are the other post-identification feedback questions in the sense that they involve specific details and do not have a strong evaluative undertone. And, interestingly, two of these three judgments (length of view and distance) are the only two of all the standard post-identification feedback judgments that have been reliably shown to be unaffected by post-identification feedback (see meta-analysis by Steblay et al., 2014). To date, nobody has advanced a clear theoretical interpretation as to why these two judgments are not influenced by confirming feedback.

Results of the current work might help provide an explanation as to why some judgments, but not others, are influenced by post-identification feedback. Specifically, feedback might not influence witnesses' retrospective judgments when the judgments involve objective recollections of specific details of the witnessed event. In these cases, witnesses might be relying more on their actual memory of the event (i.e., an internal cue) than they do when they are making evaluative, meta-cognitive judgments. Such is the case for questions about the time that the culprit was in view and how far away the culprit was during witnessing. But what about witnesses' judgments of how long it took them to make an identification, which also refers to an objective detail from the witness' memory yet has been shown to be influenced by confirming

feedback? Perhaps witnesses do not encode clear memories for how long it took them to make an identification. After all, unlike their memories of the witnessed event itself, the time involved in making an identification lacks temporal development markers or scene changes that would allow witnesses to imagine the timeline of the event in their mind's eye. This is merely speculation, but the possibility that objective (as opposed to evaluative) judgments about specific event details might be less susceptible to distortion from feedback introduces an interesting explanation for why some judgments have shown consistent immunity to the distorting effects of feedback. Moreover, it provides a possible explanation as to why the current work failed to show an effect of confirming feedback on evaluators' judgments on the witnesses' testimonies: namely, that many of the questions asked in the testimony interview were questions that are not distorted by feedback and hence provide cues that were more diagnostic of the accuracy of the eyewitnesses' identifications.

Does Collecting Pre-Feedback Statements Protect Against Feedback Effects on Testimony?

The protective hypothesis in the current work stated that collecting pre-feedback statements from witnesses would help protect witnesses from the distorting effects of post-identification feedback. This hypothesis was derived from the cue-accessibility interpretation of the post-identification feedback effect, which postulates that witnesses rely on feedback to infer testimony-relevant judgments to the extent that internal cues to these judgments are weak (Charman & Wells, 2012; Wells & Bradfield, 1998, 1999). Asking witnesses questions about their certainty, view, attention, etc., before giving them feedback was expected to create and/or strengthen witnesses' internal cues to these judgments, thereby decreasing their reliance on feedback when they answered such questions during testimony.

Unfortunately, due to the failure of the current work to replicate the feedback-impairment effect on evaluators' judgments of eyewitnesses in the control (no pre-feedback statements) condition, I was unable to conduct a statistical test of the protective hypothesis. Despite being unable to adequately test this hypothesis, inspecting the pattern of evaluators' belief rates in the pre-feedback statements collected-but-not-shown condition can provide some potentially useful information regarding the extent to which collecting pre-feedback statements serves as a protectant against feedback-induced distortion.

Evaluators who viewed witnesses who had provided pre-feedback statements were just as able to discriminate between accurate and mistaken eyewitnesses when the witnesses received feedback as when they did not. This finding indicates that feedback did not impair evaluators' abilities to discriminate between accurate and mistaken eyewitnesses in the pre-feedback statements collected-but-not-shown condition. The possibility exists, then, that collecting pre-feedback statements did in fact buffer witnesses from being influenced by post-identification feedback. But that argument can be made only if the feedback-impairment effect had been observed in the control condition. Because it was not, these data are unable to statistically show a discriminability improvement among witnesses who gave pre-feedback statements.

Although the results cannot be directly compared, it is instructive to examine differences in the patterns of evaluators' belief rates from the original Smalarz and Wells (2014) study (in which feedback impaired evaluators' abilities to discriminate between accurate and mistaken eyewitnesses) and the pre-feedback statements collected-but-not-shown condition of the current study (in which feedback did not impair discrimination). Specifically, the reduction in evaluators' abilities to discriminate in the original Smalarz and Wells study was due to inflation in the believability of mistaken eyewitnesses without corresponding inflation in the believability

of accurate eyewitnesses. In the current study in which witnesses provided pre-feedback statements, feedback inflated the believability of mistaken eyewitnesses, but it also inflated the believability of accurate eyewitnesses. As a result, evaluators were more likely to believe eyewitnesses who received feedback but they were still able to discern to some degree which witnesses were accurate and which were mistaken.

This finding is inconsistent with the cue-accessibility conceptualization of the post-identification feedback effect. Specifically, the cue-accessibility conceptualization postulates that accurate eyewitnesses have stronger internal memory cues regarding retrospective judgments than do mistaken eyewitnesses and are thus less strongly influenced by external cues such as post-identification feedback for inferring these judgments. In the current study, however, accurate eyewitnesses were just as influenced by the feedback as were mistaken eyewitnesses. Hence, the evaluators' abilities to maintain discrimination between accurate and mistaken eyewitnesses in this work was not due to a reduction in the influence of feedback on mistaken eyewitnesses. Rather, evaluators' abilities to discriminate between accurate and mistaken witnesses who received feedback were driven by the fact that feedback had just as strong of an effect on the accurate eyewitnesses. Although this result poses some problems for the cue-accessibility conceptualization, it would be premature to abandon the cue-accessibility conceptualization altogether. Indeed, the cue-accessibility conceptualization accounts for a number of important findings in the eyewitness literature such as the finding that feedback has a stronger influence on mistaken than on accurate eyewitnesses (see meta-analysis by Steblay et al, 2014), and the finding that prophylactic measures such as asking witnesses about their certainty or their view prior to giving them feedback helps to protect them against inflation from feedback (Wells & Bradfield 1998, 1999; Quinlivan et al., 2009). However, the current work does suggest

that there are aspects of the post-identification feedback effect that we do not yet fully understand.

Although the cue-accessibility conceptualization would not have predicted that accurate eyewitnesses would be as strongly influenced by feedback as were mistaken eyewitnesses, the conceptualization could still account for these results if, for some reason, even accurate witnesses lacked access to strong internal cues about their own accuracy. This could happen if accurate witnesses failed to experience true recognition and instead came to identify the culprit through secondary processes such as process-of-elimination or strategic guessing. For example, perhaps the culprit did not bear close resemblance to his appearance in the video, but he was still the lineup member who was the best choice relative to the other lineup members. If this were the case, even accurate eyewitnesses would still be susceptible to influence from feedback due to making their identifications in the absence of strong internal cues to accuracy. One problem with this reasoning, however, is that the current research used the same stimulus materials that were used in the original Smalarz and Wells (2014) study in which feedback influenced mistaken eyewitnesses more than it influenced accurate eyewitnesses. And yet, in the current work, evaluators' perceptions of witnesses' confidence, attention paid during witnessing, view during witnessing, and so forth, were considerably lower than they were in the original Smalarz and Wells (2014) study. For example, in the absence of feedback, evaluations of accurate eyewitnesses in the current study ranged from 4.8 (confidence) to 6.4 (accuracy of the event description) whereas evaluations of accurate eyewitnesses in the original study ranged from 6.4 (image of culprit) to 7.0 (confidence)⁵. Hence, there is some reason to believe that accurate eyewitnesses in the current study might not have had a strong internal experience of recognition

⁵ One exception in the original Smalarz and Wells (2014) study was the 'evidence sufficiency' measure, which fell at 3.9 for accurate witnesses who received no feedback.

when they identified the culprit and hence remained rather susceptible to the influence of confirming post-identification feedback.

Another possibility that could account for the fact that feedback influenced accurate eyewitnesses as much as it influenced mistaken eyewitnesses is that, for some reason, eyewitnesses in this study were skeptical about the likely accuracy of their identifications. This might have happened as a result of participant-pool contamination. For example, it is possible that participants in this study had familiarity with eyewitness research due to having taken a course at the university that exposes students to the psychology of eyewitness identification and the concept of mistaken identification. To the extent that witnesses in this study were familiar with the idea that the actual culprit is often not in the lineup, then even accurate eyewitnesses might have made their identifications while still doubting their likely accuracy. As a result, these witnesses might have been more susceptible to the influence of feedback than they would have been if they had no reason to doubt the validity of their recognition experience. However, there is no reason to believe that the semesters during which data were collected for the current study were unique in terms of student learning about eyewitness identification. Moreover, procedures were utilized in this work to minimize the chances that students with knowledge of eyewitness identification would participate (i.e., restricting sign-ups to students in classes that had not covered eyewitness identification) and all participants were probed for their knowledge about eyewitness identification during debriefing. Hence, it is rather unlikely that systemic participant pool contamination is responsible for the unusual pattern of results observed in the current work.

Why No Correction Effects from Showing Evaluators Witnesses' Pre-feedback Statements?

The corrective hypothesis in the current work predicted that providing evaluators with a record of witnesses' pre-feedback statements would help them better discriminate between accurate and mistaken eyewitnesses who received feedback. Unfortunately, the results did not support this hypothesis. Indeed, the pattern of evaluators' belief rates was consistent across the statements collected-but-not-shown condition and the statements collected-and-shown condition, indicating that showing witnesses' pre-feedback statements to evaluators did not improve their abilities to discern accurate from mistaken eyewitness testimony. There are three possible reasons as to why this was.

The first potential explanation for why showing witnesses' pre-feedback statements to evaluators did not improve evaluators' abilities to discriminate between accurate and mistaken testimony is that the testimony of witnesses who provided pre-feedback statements was not distorted by the post-identification feedback. This goes to the heart of the theory underlying the protective hypothesis: namely, that asking witnesses questions about their confidence and the quality of the witnessing conditions protects them from the distorting effects of feedback. To the extent that collecting pre-feedback statements kept the witnesses' testimonies from being influenced by the feedback, then showing evaluators witnesses' pre-feedback statements would not improve discrimination. After all, pre-feedback eyewitness statements are only useful to the extent that they provide information that is more diagnostic of identification accuracy than are the eyewitnesses' testimonies. However, this interpretation begs the question as to why feedback had a main effect on evaluators' perceptions of the eyewitnesses. Evaluators were more likely to believe witnesses who received confirming feedback following their identification, regardless of whether they made an accurate or mistaken identification. Hence, it appears that collecting pre-

feedback statements from eyewitnesses did not completely protect them from being influenced by the feedback; rather, as discussed above, the feedback simply inflated the testimony of accurate eyewitnesses as much as it inflated the testimony of mistaken eyewitnesses.

Another possible reason why showing witnesses' pre-feedback statements to evaluators did not have a corrective effect on evaluators' judgments of the eyewitnesses has to do with the mediums in which the testimony interview and pre-feedback statements were presented. Specifically, witnesses' testimonies were filmed with a high-quality camera and showed a clear view of the witnesses during testimony. The sound quality was excellent, and witnesses faced the camera as they responded to the interviewer's questions. In contrast, the pre-feedback statements were recorded using a hidden camera that was placed on top of a shelf approximately 6 feet away from the witnesses. It captured both the witness and the lineup administrator, and the video and sound quality were slightly impoverished. Hence, it is possible that the vividness of witnesses' testimonies relative to the pre-feedback statements made the testimonies particularly influential on evaluators' judgments. Additionally, evaluators always viewed the pre-feedback statements after they viewed the testimony interviews. Perhaps evaluators formed relatively strong impressions of the eyewitnesses before they viewed the pre-feedback statements and then did not sufficiently adjust their perceptions based on new information gleaned from the pre-feedback statements.

A final possible reason why showing witnesses' pre-feedback statements to evaluators did not improve discrimination could be because evaluators either did not notice inconsistencies between the witnesses' testimonies and their pre-feedback statements or they were unsure about how to interpret inconsistencies. Although evaluators were given instructions suggesting that the witnesses' testimonies might have been bolstered and were therefore potentially unreliable (i.e.,

“It has been argued that statements taken immediately (at the time of the identification) might be better for purposes of evaluating the reliability of an eyewitness than are statements taken at a later time (such as during testimony) because of the possibility that other events might have occurred that bolstered the eyewitness's testimony.”), evaluators might not have known exactly if, how, or why the witnesses’ testimonies were bolstered. In a real trial, there would be an opportunity for the defense to cross-examine a witness regarding inconsistencies between statements made at the time of the identification and testimony during trial. A defense attorney could point out that the witness’ certainty statement became inflated, for example, and ask the witness to account for why the inflation occurred. Indeed, much of the research showing that jurors are sensitive to inconsistencies in witnesses’ testimonies has used cross-examination to uncover and bring attention to the inconsistencies (e.g., Berman & Cutler, 1996; Berman, Narby, & Cutler, 1995). Bradfield and McQuiston (2004) and Jones, Williams, and Brewer (2008), however, found that mock-jurors perceptions were more favorable to the defense even when inconsistencies between the witness’ initial confidence and later confidence were not explicitly challenged by the defense. In these studies, however, the inconsistencies were rather salient; at trial, the witness reported that she was positive and she reported that she had been “not sure” at the time of the identification. In the current work, evaluators watched witnesses answer a number of different questions during pre-feedback statement questioning. Perhaps it was difficult for evaluators to keep track of what witnesses had said at the time of testimony versus during pre-feedback statement questioning. Moreover, to the extent that evaluators did recognize inconsistencies, they might have been uncertain about what to make of the inconsistencies. Drawing their attention to the inconsistencies and pressing witnesses on the cause of the

inconsistencies might have been necessary to elicit skepticism among evaluators who viewed witnesses' pre-feedback statements.

CHAPTER 6

CONCLUSION

This research sought to test a novel safeguard for protecting against and correcting for the impairing effects of confirming post-identification feedback on evaluators' abilities to discriminate between accurate and mistaken eyewitness testimony. The value of such a safeguard is readily apparent in a legal system that has wrongfully convicted more than 220 individuals on the basis of highly confident but mistaken eyewitness identification testimony (Innocence Project, 2015; Smalarz & Wells, 2015). Unfortunately, due to observing rather unexpected data patterns in the current work, this research was unable to provide conclusive evidence regarding the efficacy of the pre-feedback eyewitness statements safeguard. Importantly, however, this research is rich in its theoretical value because it sparks a number of interesting questions that have the potential to further develop our understanding of how post-identification feedback influences eyewitnesses.

The most promising avenue for future research involves pursuing a better understanding of the conditions under which feedback does not impair the abilities of evaluators to discriminate between accurate and mistaken testimony. In the control (no pre-feedback statements) condition of this research, feedback did not impair evaluators' abilities to discriminate between accurate and mistaken eyewitnesses. Not only that, but feedback did not inflate the belief of the eyewitnesses at all. Despite a considerable body of work investigating potential moderators of the post-identification feedback effect (e.g., Charman et al., 2010; Lampinen et al., 2007; Neuschatz et al., 2007; Quinlivan et al., 2009; Quinlivan et al., 2010; Wells & Bradfield, 1998, 1999), no remedy has been successful at eliminating the effects of feedback on eyewitnesses.

Instead, prophylactic and remedial solutions only reduced the impact of the feedback on eyewitnesses. To the extent that the current work has revealed a mechanism for eliminating the post-identification feedback effect, it bears important implications for both theory and practice.

How might future research test the conditions under which feedback does not influence evaluations of eyewitnesses? The current work proposed two potential reasons for the null effects of feedback on evaluators' judgments. First, witnesses in this study did not answer questions about their certainty, attention, view, and so forth following the feedback but prior to the testimony interview. Second, the testimony interview in the current study was much longer and more elaborate than the testimony interview in the original Smalarz and Wells (2014) study. Future research might manipulate these variables to examine whether they moderate the effects of feedback on evaluator's judgments of eyewitness identification testimony. For example, witnesses could make an identification and receive confirming feedback or no feedback following their identification. Then, the witnesses could either answer or not answer the standard post-identification feedback questions regarding their identification and the witnessing experience. Witnesses could then go on to provide testimony in either a short or a long testimony interview, and these testimony tapes could be later shown to evaluators. This type of research design could test whether asking witnesses questions immediately following feedback and prior to the witnesses giving testimony influences their testimony and also whether the type of testimony interview (short versus long) has implications for evaluators' belief of eyewitnesses. Such research could shed light on heretofore undiscovered aspects of the post-identification feedback effect.

Eyewitness identification research has developed considerably since Justice Brennan first warned against the dangers of overly-persuasive eyewitnesses, but concerns about the

incriminating power of mistaken eyewitness testimony still persist. Today, however, we know much more about how mistaken identifications can end up resulting in wrongful convictions. The answer lies partly in the fact that the credibility of eyewitness identification testimony is driven largely by eyewitnesses' self-reports about how confident they were in their identification and how favorable the witnessing conditions were at the time of the event. As the research has shown time and again, these eyewitness self-reports are highly malleable and susceptible to distortion from all-too-common extraneous influences such as confirming post-identification feedback. Because it is impossible to fully eliminate post-identification feedback and other post-identification factors that influence eyewitnesses, there is great value in continued attempts to mitigate the effects of these contaminants on eyewitness identification evidence. The trust placed by the legal system in eyewitness identification evidence for solving crimes demands that the credibility of an eyewitness be determined by the witness' memory, not by extraneous information provided to the witness by the legal system.

REFERENCES

- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 6, 1-62.
- Berman, G. L., & Cutler, B. L. (1996). Effects of inconsistencies in eyewitness testimony on mock-juror decision making. *Journal of Applied Psychology*, 81(2), 170–177.
doi:10.1037//0021-9010.81.2.170
- Berman, G. L., Narby, D. J., & Cutler, B. L. (1995). Effects of inconsistent eyewitness statements on mock-jurors' evaluations of the eyewitness, perceptions of defendant culpability and verdicts. *Law and Human Behavior*, 19(1), 79–88.
doi:10.1007/BF01499074
- Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, K. E. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: A meta-analysis of two variables associated with initial memory strength. *Psychology, Crime, & Law*, 18, 473-490. doi: 10.1080/1068316X.2010.508458
- Bothwell, R.K., Deffenbacher, K.A., & Brigham, J.C. (1987). Correlation of eyewitness accuracy and certainty: Optimality hypothesis revisited. *Journal of Applied Psychology*, 72, 691-695. doi:10.1037/0021-9010.72.4.691.
- Bradfield, A. & McQuiston, D. E. (2004). When does evidence of eyewitness confidence inflation affect judgments in a criminal trial? *Law and Human Behavior*, 28, 369-387.
doi:10.1023/B:LAHU.0000039331.54147.ff.

- Bradfield, A. L. & Wells, G. L. (2000). The perceived validity of eyewitness identification testimony: A test of the five Biggers criteria. *Law and Human Behavior, 24*, 581-594.
doi:10.1023/A:1005523129437
- Bradfield, A. L., Wells, G.L, & Olson, E.A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology, 87*, 112-120. doi:10.1037/0021-9010.87.1.112
- Brewer, N., Potter, R., Fisher, R. P., Bond, N., & Luszcz, M. A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Applied Cognitive Psychology, 13*, 297– 313.
doi:10.1002/(SICI)1099-0720(199908)13:4<3C297::AID-ACP578%3E3.3.CO;2-J
- Brigham, J. C., & Bothwell, R. K. (1983). The ability of prospective jurors to estimate the accuracy of eyewitness identification. *Law and Human Behavior, 7*, 19-30.
doi:10.1007/BF01045284
- Charman, S. D., Carlucci, M., Vallano, J., & Gregory, A. (2010). The selective cue integration framework: A theory of post-identification witness confidence assessment. *Journal of Experimental Psychology: Applied, 16*, 204–218. doi:10.1037/a0019495
- Charman, S. D., & Wells, G. L. (2012). The moderating effect of ephoric experience on post-identification feedback: A critical test of the cues-based inference conceptualization. *Applied Cognitive Psychology, 26*, 243-250. doi: 10.1002/acp.1815
- Cialdini, R. B. & Trost M. R. (1998). Social influence: social norms, conformity, and compliance. In *The Handbook of Social Psychology*, ed. DT Gilbert, ST Fiske, G Lindzey, 2:151–92. Boston: McGraw-Hill. 4th ed.

- Clifford, B. R. and Scott, J. (1978). Individual and situational factors in eyewitness testimony. *Journal of Applied Psychology*, 63, 352-359. doi: 10.1037//0021-9010.63.3.352
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.
- Cutler, B.L., Penrod, S.D., & Dexter, H.R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, 14, 185-191. doi: 10.1007/BF01062972
- Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision-making in eyewitness identification cases. *Law and Human Behavior*, 12, 41-55. doi: 10.1007/BF01064273
- Deffenbacher, K. A., Brown, E. L., & Sturgill, W. (1978). Some predictors of eyewitness memory accuracy. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press.
- Deffenbacher, K. A., & Loftus, E. E (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law and Human Behavior*, 6, 15-30.
doi:10.1007/BF01049310
- Douglass, A. B., Neuschatz, J. S., Imrich, J., & Wilkinson, M. (2010). Does post-identification feedback affect evaluations of eyewitness testimony and identification procedures? *Law and Human Behavior*. doi:10.1007/s10979-009-9189-5
- Dysart, J. E. & Lindsay, R. C. L. (2006). Show-up identifications: Suggestive technique or reliable method? In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toggia (Eds.), *Handbook of eyewitness psychology: Memory for people* (pp. 557–572). Mahwah, NJ: Erlbaum.
- Ebbinghaus, H. E. (1885). *Memory: A contribution to experimental psychology*. New York: Dover, 1964.

- Fisher, R. P., & Craik, F. I. (1977). Interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Human Learning and Memory*, 3 (6), 701-711.
doi:10.1037//0278-7393.3.6.701
- Fox, S. G., & Walters, H. A. (1986). The impact of general versus specific expert testimony and eyewitness certainty upon mock juror judgment. *Law and Human Behavior*, 10, 215-228.
doi:10.1007/BF01046211
- Garrioch, L. & Brimacombe, C. A. E. (2001). Lineup administrators' expectations: Their impact on eyewitness confidence. *Law and Human Behavior*, 25, 299-315.
doi:10.1023/A:1010750028643
- Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1986). Enhancement of eyewitness memory with the Cognitive Interview. *American Journal of Psychology*, 99, 385-401. doi:10.2307/1422492
- Greathouse, S. M., & Kovera, M. B. (2009). Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification. *Law and Human Behavior*, 33, 70-82. doi:10.1007/s10979-008-9136-x
- Greene v. McElroy*, 360 U.S. 474, 1959.
- Hafstad, G. S., Memon, A., & Logie, R. (2004). Post-identification feedback, confidence and recollections of witnessing conditions in child witnesses. *Applied Cognitive Psychology*, 18, 901-912. doi:10.1002/acp.1037
- Innocence Project (2015). www.innocenceproject.org. Retrieved March 13, 2015.
- Jones, E. E., Williams, K. D., & Brewer, N. (2008). "I had a confidence epiphany!": Obstacles to combating post-identification confidence inflation. *Law & Human Behavior*, 32, 164-176. doi:10.1007/s10979-007-9101-0.

- Kassin, S. M., Ellsworth, P., & Smith, V.L. (1989). The "general acceptance" of psychological research on eyewitness testimony. *American Psychologist*, 44, 1089-1098.
doi:10.1037/0003-066X.44.8.1089.
- Krafka, C, & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality and Social Psychology*, 49, 58-69.
doi:10.1037//0022-3514.49.1.58
- Lampinen, J., Scott, J., Pratt, D., Leding, J. K., & Arnal, J. D. (2007). "Good, you identified the suspect. . .but please ignore this feedback." Can warnings eliminate the effects of post-identification feedback? *Applied Cognitive Psychology*, 21, 1037–1056.
doi:10.1002/acp.1313
- 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.
doi:10.1037//0021-9010.63.3.345
- Lindsay, R. C. L., Wells, G. L., & Rumpel, C. (1981). Can people detect eyewitness identification accuracy within and across situations? *Journal of Applied Psychology*, 66, 79-89.
doi:10.1037//0021-9010.66.1.79
- Loftus, E. F. (1974). Reconstructing memory: The incredible eyewitness. *Psychology Today*, 8, 116-119.
- Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
doi:10.1037//0021-9010.66.1.79
- Narby, D. J., & Cutler, B. L. (1994). Effectiveness of voir dire as a safeguard in eyewitness cases. *Journal of Applied Psychology*, 79, 724-729. doi: 10.1037/0021-9010.79.5.724

National Institute of Justice. Technical Working Group for Eyewitness Evidence (1999).

Eyewitness evidence: A guide for law enforcement. Washington, DC: United States Department of Justice, Office of Justice Programs.

Neuschatz, J. S., Lawson, D. S., Fairless, A. H., Powers, R. A., Neuschatz, J. S., Goodsell, C. A., & Togli, M. P. (2007). The mitigating effects of suspicion on post-identification feedback and on retrospective eyewitness memory. *Law and Human Behavior, 31*, 231-247. doi:10.1007/s10979-006-9047-7

Neuschatz, J. S., Preston, E. L., Burkett, A. D., Togli, M. R., Lampinen, J. M., Neuschatz, J. S., Fairless, A. H., Lawson, D. S., Powers, R. A., & Goodsell, C. A. (2005). The effects of post-identification feedback and age on retrospective eyewitness memory. *Applied Cognitive Psychology, 19*, 435-453. doi:10.1002/acp.1084

Police Executive Research Forum (2013). *A National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies*. Report submitted to the National Institute of Justice, March 8, 2013.

Quinlivan, D. S., Neuschatz, J. S., Jimenez, A., Cling, A. D., Douglass, A. B., & Goodsell, C. A. (2009). Do prophylactics prevent inflation? Post-identification feedback and the effectiveness of procedures to protect against confidence-inflation in earwitnesses. *Law & Human Behavior, 33*, 111-121. doi: 10.1007/s10979-008-9132-1

Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of post-identification feedback on eyewitness identification and nonidentification confidence. *Journal of Applied Psychology, 89*, 334-346. doi:10.1037/0021-9010.89.2.334

Shaw, J. S. III (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied, 2*, 126-146. doi: 10.1037/1076-898X.2.2.126

- Shaw, J. S., III & McClure, K. A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior, 20*, 629-654.
doi:10.1007/BF01499235
- Smalarz, L., Norris, S., & Wells, G. L. (2014). *The legal engine that couldn't: Cross-examination fails to correct for feedback effects on evaluations of eyewitness testimony*. Paper presented at the annual American Psychology-Law Society Conference, New Orleans, LA.
- Smalarz, L. & Wells, G. L. (2013). Eyewitness certainty as a system variable. (pp. 161-177) In B. L. Cutler (Ed.), *Reform of eyewitness identification procedures*. Washington DC: American Psychological Association Press.
- Smalarz, L. & Wells, G. L. (2014). Post-identification feedback to eyewitnesses impairs evaluators' abilities to discriminate between accurate and mistaken testimony. *Law and Human Behavior, 38*, 194-202. doi: 10.1037/lhb0000067
- Smalarz, L., & Wells, G. L. (2015). Contamination of eyewitness self-reports and the mistaken identification problem. *Current Directions in Psychological Science, 24*(2) 120–124.
- Sporer, S. L. Penrod, S. Read, D. & Cutler, B. L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*, 315-327. doi:10.1037/0033-2909.118.3.315
- Stebly, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21*, 283-298.
doi:10.1023/A:1024890732059.

- Stebly, N. M., Wells, G. L., & Douglass, A. L. (2014). The eyewitness post-identification feedback effect 15 years later: Theoretical and policy implications. *Psychology, Public Policy, and Law*, 20, 1–18. doi:10.1037/law0000001
- Stinson, Y., Devenport, J. L., Cutler, B. L., & Kravitz, D. A. (1996). How effective is the presence-of-counsel safeguard? Attorney perceptions of suggestiveness, fairness, and correctability of biased lineup procedures. *Journal of Applied Psychology*, 81, 64-75. doi:10.1037//0021-9010.81.1.64
- Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373. doi: 10.1037/h0020071
- United States v. Ash*, 413 U.S. 300, 1973.
- Van Wallendael, L. R., Cutler, B. L., Devenport, J., & Penrod, S. (2006). Mistaken identification = Erroneous conviction? Assessing and improving legal safeguards. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (pp. 557–572). Mahwah, NJ: Erlbaum.
- Walters, C. M. (1985). Admission of expert testimony on eyewitness identification. *California Law Review*, 73, 1402–1430.
- Wells, G. L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36, 1546-1557. doi:10.1037/0022-3514.36.12.1546.
- 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, 360-376. doi:10.1037/0021-9010.83.3.360

- Wells, G. L. & Bradfield, A. L. (1999). Distortions in eyewitnesses' recollections: Can the post-identification feedback effect be moderated? *Psychological Science, 10*, 138-144.
doi:10.1111/1467-9280.00121
- Wells, G. L., Ferguson, T. J., & Lindsay, R. C. L. (1981). The tractability of eyewitness confidence and its implications for triers of fact. *Journal of Applied Psychology, 66*, 688-696. doi:10.1037//0021-9010.66.6.688
- Wells, G. L., Greathouse, S. M., & Smalarz, L. (2011). Why do motions to suppress suggestive eyewitness identifications fail? In B.L. Cutler (Ed.), *Conviction of the Innocent: Lessons from Psychological Research*. Washington DC: American Psychological Association Press.
- Wells, G. L., & Luus, E. (1990). Police lineups as experiments: Social methodology as a framework for properly-conducted lineups. *Personality and Social Psychology Bulletin, 16*, 106-117. doi:10.1177/0146167290161008
- Wells, G. L., Olson, E., & Charman, S. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied, 9*, 42-52.
doi:10.1037/1076-898X.9.1.42
- Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior, 33*, 1-24. doi:10.1007/s10979-008-9130-3
- 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*, 603-647. doi:10.1023/A:1025750605807

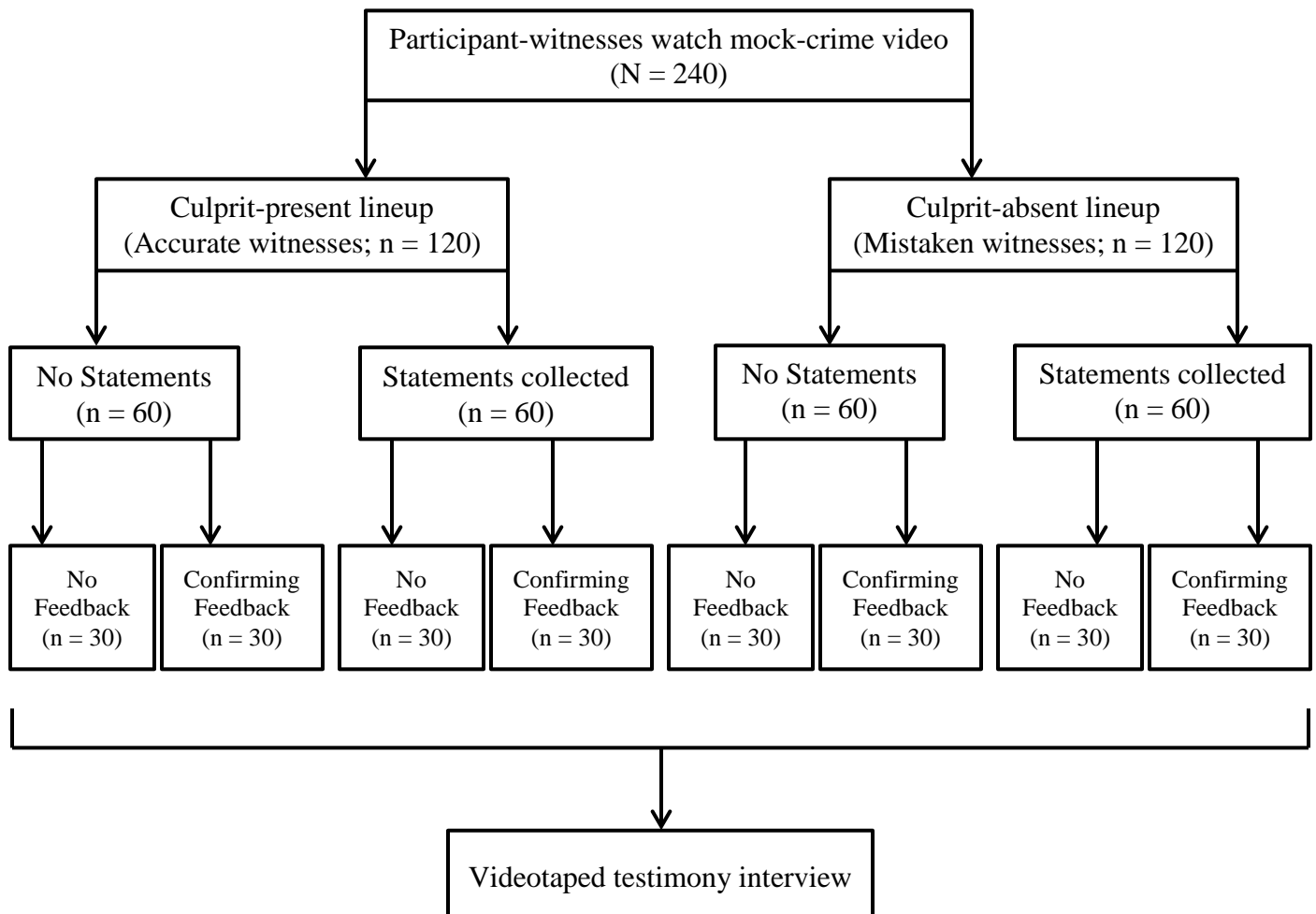
- Wigmore, J. H. (1970). *Evidence in trials at common law. (Chadbourn Revision)*. Boston: Little, Brown, and Company.
- Wilford, M. M. & Wells, G. L. (2013). Eyewitness system variables. In B. L. Cutler (Ed.), *Reform of eyewitness identification procedures*. Washington DC: American Psychological Association Press.
- Wright, D. B. & Skagerberg, E. M. (2007). Post-identification feedback affects real eyewitnesses. *Psychological Science*, *18*, 172-178. doi:10.1111/j.1467-9280.2007.01868.x
- Yarmey, A. D., & Jones, H. P. T. (1983). Is the psychology of eyewitness identification a matter of common sense? In S. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 13--40). Chichester, England: Wiley.

APPENDIX A. EXPERIMENTAL DESIGN FOR THE WITNESS PHASE

Experimental Design for the Witness Phase

	Statements Not Collected		Statements Collected	
	No Feedback	Confirming Feedback	No Feedback	Confirming Feedback
Mistaken	$n=30$	$n=30$	$n=30$	$n=30$
Accurate	$n=30$	$n=30$	$n=30$	$n=30$

Note. 2 (Pre-feedback statements: collected vs. not collected) x 2 (Identification accuracy: accurate vs. mistaken) x 2 (Post-identification feedback: no feedback vs. confirming feedback) between-subjects design.

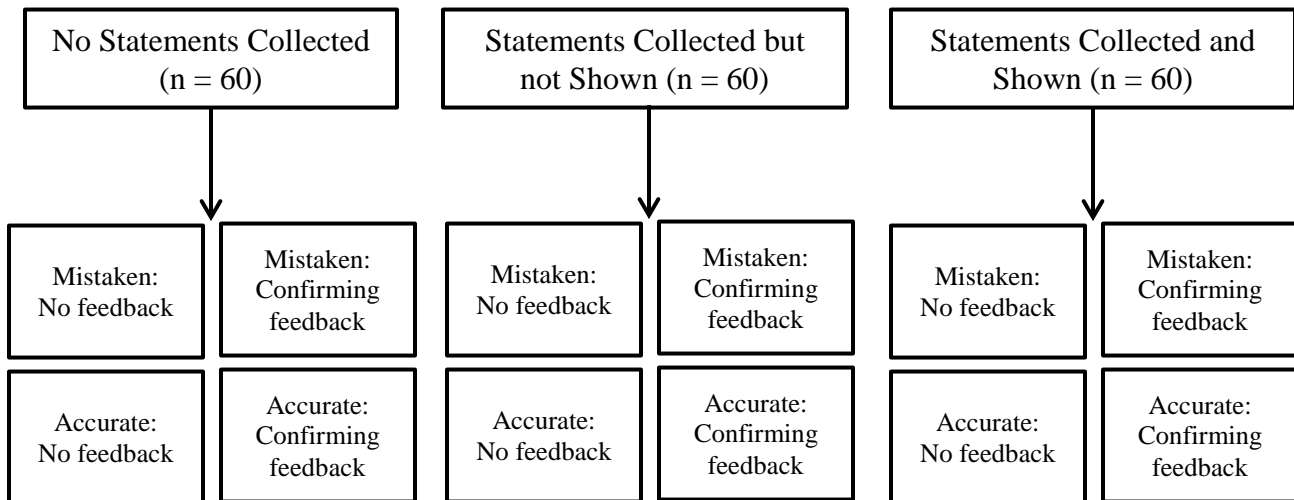
Experimental Procedure for the Witness Phase

APPENDIX B. EXPERIMENTAL DESIGN FOR THE EVALUATOR PHASE

Experimental Design for the Evaluator Phase

Identification Accuracy	No Statements Collected		Statements Collected but not Shown		Statements Collected and Shown	
	No Feedback	Confirming Feedback	No Feedback	Confirming Feedback	No Feedback	Confirming Feedback
Mistaken	$n=60$	$n=60$	$n=60$	$n=60$	$n=60$	$n=60$
Accurate	$n=60$	$n=60$	$n=60$	$n=60$	$n=60$	$n=60$

Note. 3 (Pre-feedback statements: not collected vs. collected but not shown vs. collected and shown) x 2 (Identification accuracy: accurate vs. mistaken) x 2 (Post-identification feedback: no feedback vs. confirming feedback) mixed design.

Experimental Procedure for the Evaluator Phase

APPENDIX C. STILL SHOTS FROM THE WITNESSED EVENT



APPENDIX D. CULPRIT-ABSENT LINEUP

1



2



3



4



5



6

APPENDIX E. CULPRIT-PRESENT LINEUP

1



2



3



4



5



6

Note. Culprit is number two.

APPENDIX F. PRE-FEEDBACK STATEMENT QUESTIONS

	Question
<i>Certainty</i>	How sure are you that you identified the right person from the lineup?
<i>Basis</i>	So you felt as though you got a good enough look at the person to make an identification from the lineup?
<i>Ease</i>	How easily were you able to pick out the person from the lineup?
<i>Time</i>	About how long did it take you to pick him out?
<i>View</i>	At the time you witnessed this, how well were you able to see or observe the events taking place?
<i>Details</i>	How well were you able to see the facial details of the person who switched the bags?
<i>Attention</i>	How closely were you focusing on the person's face?
<i>Image</i>	In your mind's eye, how well can you still see that person?
<i>Willing</i>	On the basis of your memory of the person, how willing would be to testify in court that the person you identified was person who switched the bags?
<i>Strangers</i>	Do you consider yourself to be a person who has a pretty good memory for people?

APPENDIX G. TESTIMONY INTERVIEW SCRIPT

Today's date is _____. Thank you for being here today.

Do you understand that your testimony today concerns something that you observed in the Des Moines International airport on *today's date*?

Do you understand that an individual has been accused of planting a bomb in a suitcase and switching that suitcase with another passenger at the Des Moines airport?

And you had an occasion, then, to be able to observe an individual who switched a suitcase at the Des Moines airport? Wait for "yes". That's what we're going to be talking about today.

Had you ever been to that airport before?

You had an occasion to view a line of people waiting to check-in at one of the airline counters, is that right?

How many people were in that line?

Please tell me in as much detail as possible exactly what it was that you witnessed at the check-in counter.

Did any of the other passengers in line seem to notice the bag-switching?

Can you tell me anything about the bag?

The person who you saw switch the bags, can you describe this person to me?

What made this person stand out to you?

Was there anything in particular about this person's hair, or face, or clothing that made him stand out to you?

While you were observing the check-in line, how much attention were you paying to the face of the person who you saw switch the bags?

And how good of a view would you say you got of the person who switched the bags?

How well were you able to make out specific features of the face of the person who switched the bags?

How clear is the image you have in your memory of the person you saw who switched the bags?

You viewed a photo spread containing various individuals who fit the description of the person who switched the bags, is that right?

How many photos were there?

And you identified someone as the person who switched the bags, is that right?

How certain were you when you made your identification that the person you identified was the person who you saw switch the bags?

To what extent do you feel that you had a good basis, or enough information, to make an identification?

How difficult would you say it was for you to determine which person in the photo spread was the person who switched the bags?

From the time you began viewing the photo spread, how long do you estimate it took you to make a decision?

Generally, how good is your recognition memory for the faces of strangers you have encountered on only one prior occasion?

How resistant were you to the idea of testifying today about what you witnessed and about your identification?

Alright—those are all the questions I have for you.

APPENDIX H. TESTIMONY JUDGMENT MEASURES

Measure	Question	Scale
<i>Belief</i>	Do you think that the witness' identification from the photo lineup of the man who switched the bags was an accurate identification or an inaccurate identification?	Accurate -or- Inaccurate
<i>Convincing</i>	How convincing was the witness?	1 (Not at all) to 10 (Totally)
<i>Confidence</i>	How confident do you believe the witness felt that the person s/he identified was the person who switched the bags?	1 (Not at all) to 10 (Totally)
<i>View</i>	How good of a view do you think the witness got of the man who switched the bags?	1 (Very poor) to 10 (Very good)
<i>Image</i>	How clear of a mental image do you think the witness had in memory of the person who switched the bags?	1 (Not at all) to 10 (Very)
<i>Attention</i>	How much attention do you think the witness was paying when s/he witnessed the event?	1 (No attention) to 10 (Total Attention)
<i>Strangers</i>	How good do you think the witness is at remembering faces of strangers?	1 (Very poor) to 10 (Very good)
<i>Culprit Description</i>	How accurate of a description do you think the witness provided about the person who switched the bags?	1 (Not at all) to 10 (Extremely)
<i>Event Description</i>	How accurate of a description do you think the witness provided about the events that took place?	1 (Not at all) to 10 (Extremely)
<i>Additional Evidence (Reverse coded as Evidence Sufficiency)</i>	If you were a juror at trial, how much additional evidence would you need to convict the person who was identified by the witness as the man who switched the bags?	1 (No additional evidence) to 10 (A lot of additional evidence)

APPENDIX I. EVALUTORS' TESTIMONY JUDGMENT MEANS

		No Pre-feedback Statements		Pre-feedback Statements Collected but not Shown		Pre-feedback Statements Collected and Shown	
		No feedback	Confirming Feedback	No feedback	Confirming Feedback	No feedback	Confirming Feedback
Mistaken	Convincing	5.1 (2.4)	5.0 (2.6)	4.5 (2.1)	5.9 (2.3)	4.5 (2.2)	5.4 (2.4)
	Confident	3.8 (1.5)	3.7 (1.7)	3.5 (1.7)	4.3 (1.7)	3.3 (1.6)	4.2 (1.7)
	Attention	5.1 (2.3)	4.7 (2.4)	4.8 (2.3)	5.8 (2.6)	5.1 (2.0)	5.6 (2.3)
	View	5.2 (2.5)	5.2 (2.3)	4.6 (2.1)	5.9 (2.2)	4.9 (2.3)	5.7 (1.8)
	Culprit Description	5.2 (2.3)	5.2 (2.1)	4.8 (2.1)	5.4 (2.1)	4.9 (1.8)	5.4 (2.1)
	Event Description	5.9 (2.2)	5.7 (2.1)	5.6 (2.3)	6.1 (2.0)	5.7 (2.0)	6.0 (2.1)
	Image	4.8 (2.5)	4.8 (2.3)	3.8 (2.0)	5.5 (2.4)	4.2 (2.1)	4.9 (2.0)
	Strangers	5.0 (2.4)	4.4 (2.2)	3.8 (1.8)	5.6 (2.4)	4.3 (1.9)	5.1 (2.1)
Evidence Sufficiency		4.1 (2.5)	3.9 (2.2)	3.4 (2.1)	3.9 (2.0)	3.6 (1.9)	3.8 (2.0)
Accurate	Convincing	5.9 (2.2)	6.5 (2.4)	5.1 (2.5)	6.4 (2.2)	5.5 (2.4)	5.9 (2.3)
	Confident	4.8 (1.6)	5.0 (1.7)	4.1 (1.8)	4.3 (1.7)	3.8 (1.6)	4.3 (1.9)
	Attention	6.1 (2.2)	7.0 (2.0)	5.1 (2.4)	6.5 (2.1)	5.4 (2.0)	6.0 (2.3)
	View	6.2 (2.2)	6.8 (2.1)	5.2 (2.3)	6.6 (2.0)	5.7 (2.1)	6.0 (2.1)
	Culprit Description	6.0 (2.0)	6.3 (1.9)	5.2 (2.3)	6.4 (2.0)	5.2 (2.2)	5.8 (2.0)
	Event Description	6.4 (2.0)	6.8 (2.0)	5.7 (2.2)	6.4 (2.0)	5.9 (2.2)	6.2 (1.8)
	Image	6.2 (2.2)	6.6 (2.2)	4.9 (2.4)	5.9 (2.2)	5.0 (2.2)	5.9 (2.3)
	Strangers	6.0 (2.1)	6.4 (2.2)	5.0 (2.6)	5.9 (2.5)	5.0 (2.3)	5.8 (2.6)
Evidence Sufficiency		5.1 (2.4)	5.7 (2.4)	4.1 (2.3)	4.7 (2.4)	4.0 (2.1)	4.6 (2.4)

Note. Mean (SD).