Sensitizing Jurors to Factors Influencing the Accuracy of Eyewitness Identification: Assessing the Effectiveness of the Henderson Instructions

Angela M. Jones

Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: http://academicworks.cuny.edu/gc_etds

Part of the Social Psychology Commons

Recommended Citation

http://academicworks.cuny.edu/gc_etds/994

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects (2014-Present) by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@gc.cuny.edu.
SENSITIZING JURORS TO FACTORS INFLUENCING THE ACCURACY OF EYEWITNESS IDENTIFICATION: ASSESSING THE EFFECTIVENESS OF THE HENDERSON INSTRUCTIONS

by

ANGELA M. JONES

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the Degree of Doctor of Philosophy, The City University of New York

2015
This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Steven Penrod

__________________________
Date                                 Chair of Examining Committee

Joshua Brumberg

__________________________
Date                                 Executive Officer

Saul Kassin

Maureen O’Connor

Brian Bornstein

Brian Cutler
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

SENSITIZING JURORS TO FACTORS INFLUENCING THE ACCURACY OF EYEWITNESS IDENTIFICATION: ASSESSING THE EFFECTIVENESS OF THE 
HENDERSON INSTRUCTIONS

by

Angela M. Jones

Advisor: Professor Steven D. Penrod

Recently, the New Jersey Supreme Court determined that jurors may not be able to effectively evaluate eyewitness evidence (New Jersey v. Henderson, 2011). Research generally supports this contention, finding that jurors do not take into account factors surrounding the commission of the crime and identification when determining the reliability of an identification (Devenport et al., 1997). Courts have implemented various safeguards to assist jurors in evaluating eyewitness evidence, including judicial instructions and expert testimony. The New Jersey Supreme Court proposed the use of judicial instructions and suggested their use would reduce the need for expert testimony. The current studies tested the efficacy of various forms of Henderson instructions and expert testimony. In the first study, jurors were sensitive to the quality of police practices on their own. Expert testimony resulted in skepticism by reducing convictions regardless of eyewitness identification quality. No version of Henderson instructions sensitized jurors to the quality of witnessing and identification conditions. Therefore, I conducted a follow up study to examine modifications to the Henderson instructions. The modified instructions incorporated features from the I-I-Eye instructions (Pawlenko et al., 2013), such as a
condensed format, prompts designed to draw jurors’ attention to how each eyewitness factor impacts identification accuracy, and making the instructions general in nature and not tailored to the facts of the case. I also examined whether having jurors evaluate the eyewitness evidence through the use of interrogatories would influence their verdict decision. The modified version of *Henderson* sensitized jurors to the quality of witnessing conditions compared to the original *Henderson* instructions. This effect occurred regardless of whether jurors evaluated the evidence before or after determining a verdict. These results suggest the original *Henderson* instructions are having little impact on jurors’ decisions. Thus, courts may wish to delay implementation of these instructions until further research can establish their effectiveness.
Acknowledgments

I wish to thank my committee members for the time and attention they each spent working on this dissertation: Steven Penrod, Saul Kassin, Maureen O’Connor, Brian Bornstein, and Brian Cutler. I have admired each of your contributions to the field of Psychology and Law for years. I feel beyond honored to have had the opportunity to work under your guidance. And I appreciate the feedback that has only served to improve the quality of my work.

Steve: It is because of your research that I entered graduate school. You asked me when I interviewed and many times afterward, who did I want to be? What job did I want to have? And my answer was always, I wanted to be like you. Of course one could only aspire to have half the career you have had, but I will work my hardest to achieve that goal. You have impacted my life in more ways than you could know. I could not have asked for a better mentor. Thank you.

To my family: Thank you for the support and encouragement when I decided to take this wild trip to New York and all the years since. Shayne, where would I be without you? You encouraged me to go back to graduate school. You believed in me when I did not believe in myself. And you sacrificed beyond measure so that I could succeed. I hope to always make you proud. I love you and am so excited to start this new journey together.
Table of Contents

Chapter 1: Overview of Eyewitness Identification and Juror Decision Making ..........1

Chapter 2: Judicial Decisions Concerning Eyewitness Testimony ......................4

Chapter 3: How do Jurors Reason? .................................................................11

Chapter 4: Eyewitness Identification Research .............................................14

Chapter 5: Juror Knowledge Regarding Eyewitness Identification Factors .......23

Chapter 6: The Efficacy of Traditional Safeguards .......................................26

Chapter 7: Conclusion ..................................................................................37

Chapter 8: Instructions and Expert Testimony .............................................39

Chapter 9: Pilot Study ..................................................................................41

Chapter 10: Study 1 Methods ........................................................................43

Chapter 11: Study 1 Results ..........................................................................49

Chapter 12: Study 1 Discussion ....................................................................55

Chapter 13: Study 2 Methods ........................................................................63

Chapter 14: Study 2 Results ..........................................................................68

Chapter 15: Study 2 Discussion ....................................................................73

Chapter 16: Conclusions and Future Directions ..........................................78

Appendix A: Original *Henderson* Instructions with Research Enhancement in Brackets 97

Appendix B: Lay Eyewitness Knowledge (Desmarais & Read, 2011) ..............103

Appendix C: Modified *Henderson* Instructions ............................................104

References ....................................................................................................111
List of Tables

Table 1. Study 1 Descriptives ........................................................................................................ 86
Table 2. Study 1 Correlation matrix among proposed dependent and mediator variable. 87
Table 3. Study 1 Sensitivity tests: Dependent variables as a function of safeguard type, and witnessing and identification conditions ........................................................................... 88
Table 4. Study 1 Mediation model with quality of identification conditions predicting conviction (PROCESS Model 4) ............................................................................................................ 89
Table 5. Study 1 Mediation model with expert testimony predicting conviction (PROCESS Model 4) .................................................................................................................. 90
Table 6. Study 2 Descriptives ........................................................................................................ 91
Table 7. Study 2 Correlation matrix among proposed mediators and dependent variables ................................................................................................................................. 92
Table 8. Study 2 Sensitivity tests: Dependent variables as a function of instruction type and evidence strength ........................................................................................................... 93
List of Figures

Figure 1. Mediational model displaying the effect of system variables on verdict mediated by eyewitness credibility ratings (PROCESS Model 4).........................94

Figure 2. Mediational model displaying the effect of expert testimony on verdict mediated by detective credibility ratings (PROCESS Model 4).........................95

Figure 3. Moderated Mediational model displaying the conditional effect of evidence strength on verdict mediated by belief in eyewitness (PROCESS Model 10)......96
Chapter 1: Overview of Eyewitness Identification and Juror Decision Making

Eyewitness identifications are not always reliable and memory for an event is not akin to a video recording. Among documented exonerations, erroneous eyewitness identifications have contributed to 35% of all wrongful convictions and 83% of wrongful convictions for robberies (National Registry of Exonerations, 2015). The volume of wrongful convictions suggest a full understanding of the factors influencing eyewitness identification accuracy may lie outside jurors’ common knowledge (Devenport, Cutler, & Penrod, 1997). The courts, troubled by these findings, are beginning to examine admissibility standards for questionable identifications and some are seeking ways to improve jurors’ ability to distinguish the quality of an identification (see New Jersey v. Henderson, 2011; Perry v. New Hampshire, 2012; Oregon v. Lawson, 2012).

The New Jersey Supreme Court recently addressed issues with the current standard for identification admissibility (i.e., the Manson test; see Chapter 2 below) as well as in-court procedures dealing with eyewitness evidence (New Jersey v Henderson, 2011). Concerning the latter, the New Jersey Supreme Court ruled that, in all cases involving eyewitness testimony, new instructions addressing case-specific factors based on sound psychological research must be given to help the jury determine the reliability of an identification (New Jersey Supreme Court, 2012a; see Appendix A for full instructions).

The Henderson Court recognized that a variety of factors can influence the accuracy of an eyewitness identification. Factors outside the control of the criminal justice system (i.e., estimator variables) typically surround the crime itself, such as the amount of time a witness is exposed to a perpetrator (Wells, 1978). In contrast, system factors, such as whether a suspect is presented to the eyewitness alone (i.e., showup) or in a lineup with multiple members, are within
the control of the criminal justice system (Wells, 1978). Unfortunately, surveys of jury eligible participants suggest lay knowledge concerning a variety of estimator and system variables is impaired (Desmarais & Read, 2011; see Appendix B). In addition, jurors have difficulty applying any eyewitness identification knowledge to examine specific case facts or appropriately weighing eyewitness evidence when determining a verdict (Cutler, Penrod, & Stuve, 1988; Cutler, Penrod, & Dexter, 1990; Devenport et al., 1997).

Social-cognitive theories have been offered to explain juror decision making, including the story model (Pennington & Hastie, 1992), elaboration likelihood model (ELM; Petty & Cacioppo, 1986) and heuristic-systematic model (HLM; Chaiken, 1980). These theories provide explanations for shortcomings in decision making, such as when evidence is too complex or cognitive effort is lacking. Eyewitness identification is one type of evidence that is often difficult for jurors to process. Yet, jurors are expected to evaluate the range of factors that can affect accuracy to determine whether an eyewitness identification is reliable.

If factors influencing eyewitness identification accuracy are not common knowledge, how can jurors’ assessments of eyewitness evidence be improved? The courts have suggested a number of procedural safeguards to protect against wrongful convictions, including voir dire, cross-examination, expert testimony, and judicial instructions. These safeguards are intended to improve jurors' ability to discriminate between good and poor witnessing and identification conditions (i.e., sensitivity; Cutler, Dexter, & Penrod, 1989). However, most of these safeguards have been ineffective at sensitizing jurors to factors influencing eyewitness accuracy.

Preliminary research suggests the current Henderson instructions are also ineffective (Berman, Bergold, Jones, Hui, & Penrod, 2015). It is unknown, however, whether a “research
enhanced” version of the Henderson instructions can increase jurors' sensitivity to eyewitness factors. And if so, do these instructions sensitize jurors equally to estimator and system variables? Additionally, the Henderson Court suggested these new instructions would greatly reduce the need for expert testimony. Yet, this assumption has been not subjected to empirical testing. In the current paper, I will examine recent court cases concerning eyewitness identifications, the theories underlying juror decision making, research on factors affecting eyewitness accuracy, jurors' knowledge and application of these factors, safeguards proposed by the courts to protect against wrongful convictions based on erroneous eyewitness identification, and new research testing the impact of instructions on juror decision-making. Examination of the Henderson instructions will inform courts about their effectiveness and add to the body of psychological research on safeguards against eyewitness misidentifications.
Before jurors hear eyewitness testimony, federal courts rely on the *Manson* test to determine the admissibility of eyewitness identifications (*Manson v. Brathwaite*, 1977). This test contains two prongs: (1) were suggestive police procedures used to secure identification?; and (2) if so, is the identification still reliable based on the five *Biggers* criteria (*Neil v. Biggers*, 1972)? These criteria include: (1) the opportunity of the witness to view the criminal at the time of the crime, (2) the witness' degree of attention, (3) the accuracy of his prior description of the criminal, (4) the witness’ certainty, and (5) the time between the crime and the identification.

The bulk of research conducted since the *Neil v. Biggers* case in 1972 has established the inadequacy of the *Manson* criteria as indicators of eyewitness reliability. For example, an eyewitness' certainty in their identification is a weak predictor of accuracy (Sporer et al., 1995). Furthermore, suggestive procedures can influence three of the five *Biggers* criteria. Specifically, giving eyewitnesses confirmatory feedback after they have made an identification, showing an eyewitness a lineup where the suspect stands out, and failing to inform eyewitnesses that a suspect may or may not be present in a lineup (i.e., biased lineup instructions) all inflate a witness' self-reported certainty, view, and attention (Steblay, Wells, & Douglass, 2014; Wells & Quinlivan, 2009).

**Perry v. New Hampshire (2012)**

The *Manson* test mandates a finding of suggestive police procedures before any additional factors are examined. Research has since established that it is not necessary for police to create suggestive circumstances for an identification to be unreliable (Shapiro & Penrod, 1986). This was precisely the issue examined in *Perry v. New Hampshire* (2012), which was the
first case involving eyewitness identification before the U.S. Supreme Court since *Manson v. Brathwaite* (1977). The Court determined “the fallibility of eyewitness evidence does not, without the taint of improper state conduct, warrant a due process rule requiring a trial court to screen such evidence for reliability before allowing the jury to assess its creditworthiness (*Perry v. New Hampshire*, 2012, p. 734).” Thus, the Court determined that trial courts were not required to review the reliability of an eyewitness identification that was made under questionable conditions outside of police action. In this case, an eyewitness made a spontaneous identification of a person surrounded by police, from over 100 feet away, in the dark, and out of a second story window. The Court determined that none of these factors violated the defendant’s due process right.

With the *Perry* decision, the Court ruled that under Federal Law an identification can only be ruled inadmissible if a defendant proves that police created suggestive circumstances surrounding an identification and the identification is unreliable based on the *Biggers* criteria. Recognizing that the jury and not the judge should determine the reliability of eyewitness testimony, the Supreme Court noted that safeguards including cross-examination and eyewitness instructions (i.e., *Telfaire*—see Chapter 6 below) were in place during Perry’s trial to sufficiently draw attention to any reliability issues (*Perry v. New Hampshire*, 2012).

**New Jersey v. Henderson (2011)**

The *Henderson* Court requested that a Special Master hear expert testimony on human memory and eyewitness identification and make recommendations for the Court to review. These recommendations included treating eyewitness evidence the same as physical trace evidence; thus requiring that the initial burden of proving identification reliability be shifted
from the defendant to the state. Second, the *Manson* test that was previously used to determine the admissibility of eyewitness identifications in pre-trial hearings should be expanded to examine all relevant system and estimator variables to determine admissibility. Lastly, the Special Master recommended that judicial instructions and expert testimony should be used to educate the jury on eyewitness factors present in a particular case.

The *Henderson* Court did not adopt all of these recommendations. In an attempt to balance defendants’ right to a fair trial and the State’s duty to maintain public safety, the Court did not agree that eyewitness evidence should be treated the same as trace evidence nor did they shift the initial burden of proving identification reliability from the defendant to the state. Similar to the decision in *Perry*, “a defendant has the initial burden of showing some evidence of suggestiveness that could lead to a mistaken identification…That evidence, in general, must be tied to a system-and not an estimator-variable (*New Jersey v. Henderson*, 2011, p. 927).” The State must provide some proof of the identification’s reliability if some suggestiveness exists, but the “ultimate burden remains on the defendant to prove a very substantial likelihood of irreparable misidentification (*New Jersey v. Henderson*, 2011, p. 927).” The Court reasoned that “courts cannot affect estimator variables; by definition, they relate to matters outside the control of law enforcement. More probing pretrial hearings about suggestive police procedures, though, can deter inappropriate police practices (*New Jersey v. Henderson*, 2011, p. 930).”

The *Henderson* Court did agree with the Special Master that the framework for evaluating eyewitness evidence should be revised to include a review of all relevant system and estimator variables, but only when a pretrial hearing is granted based upon an initial finding of suggestiveness (caused by police). The Court still anticipated that most identifications would be
ruled admissible and the jury would be responsible for determining the reliability of an identification. Acknowledging the difficulty jurors have in evaluating eyewitness evidence, the *Henderson* Court mandated new instructions addressing a variety of eyewitness identification factors that could be tailored to a specific case (see Appendix A for full instructions). New Jersey Chief Justice Stuart Rabner stated, “The [*Henderson*] instructions are designed to minimize the risk of wrongful convictions and help jurors reach informed, just decisions” (New Jersey Supreme Court, 2012b). The Court opined, “… that with enhanced jury instructions, there will be less need for expert testimony. Jury charges offer a number of advantages: they are focused and concise, authoritative (in that juries hear them from the trial judge, not a witness called by one side), and cost-free; they avoid possible confusion to jurors created by dueling experts; and they eliminate the risk of an expert invading the jury’s role or opining on an eyewitness’ credibility (*New Jersey v. Henderson*, 2011, p. 998).”

The Innocence Project suggested the new instructions should be given both before and after an eyewitness testifies (see *New Jersey v. Henderson*, 2011). However, the Court assumed that these instructions would be given at the end of the trial along with closing instructions, but suggested the “instructions may be given during trial if warranted. For example, if evidence of heightened stress emerges during important testimony, a party may ask the court to instruct the jury midtrial about that variable and its effect on memory. Trial courts retain discretion to decide when to offer instructions (*New Jersey v. Henderson*, 2011, p. 995).”

*Oregon v. Lawson* (2012)

The *Perry* and *Henderson* Courts relied upon the *Manson* test, which concerns due process violations should unreliable eyewitness identifications be admitted. In contrast, the
Classen test (used in Oregon) focuses on the rules of evidence in determining admissibility (Oregon v. Classen, 1979). Rules of evidence do not require that a defendant demonstrate an initial finding of suggestiveness by police when the reliability of an identification is questioned. Instead and in contrast to Henderson, the Oregon Supreme Court determined that “the proponents of the evidence—in identification matters, usually the state, although not necessarily so—traditionally bears the initial burden of establishing the admissibility of the proffered evidence (Oregon v. Lawson, 2012, p. 747).”

Eyewitness evidence, like all types of evidence, must be relevant, helpful to the trier of fact, and based on the witness’ own perceptions and knowledge and not that of information given to witnesses by anyone else (e.g., police giving confirmatory feedback; Oregon v. Lawson, 2012). If these requirements are satisfied by the state, the burden shifts to the defendant to establish that the “probative value is substantially outweighed by the dangers of unfair prejudice, confusion of the issues, or misleading the jury (Oregon v. Lawson, 2012, p. 742)” in order for the identification to be ruled inadmissible. Similar to Henderson, the Lawson Court suggested that the presence of estimator variables alone would not be sufficient to rule an identification inadmissible. The Court proposed that defendants rely on other safeguards, such as cross-examination, expert testimony, and eyewitness instructions to inform the jury about estimator variables.

The Lawson Court acknowledged that the criteria covered in the second prong of Manson and Classen can themselves be influenced by suggestive procedures. As a result and in departure from the Henderson Court, the Oregon Supreme Court determined that eyewitness evidence should be treated like physical trace evidence. That is, police must take care not to influence,
alter, or contaminate in any way an eyewitness’ memory. Also in stark contrast to *Henderson*, the *Lawson* Court recommended the use of expert testimony to educate jurors about factors not commonly understood to influence identification reliability and keep jurors abreast of the evolving research on eyewitness identification.

The opinion in *Henderson* suggested that the need for expert testimony would lessen with case-specific eyewitness instructions, while the *Lawson* opinion suggested expert testimony should be relied upon more. Despite past research demonstrating the ineffectiveness of eyewitness instructions (e.g., Greene, 1988), the *Henderson* Court anticipates that these new instructions will sensitize jurors to factors known to increase the risk of misidentification. However, initial findings suggest the current *Henderson* instructions are not as effective as intended (see Berman et al., 2015) and may need to be modified. In addition, no research to date has examined the relative impact of the *Henderson* instructions on jurors’ decisions compared to expert testimony.

The recent attention by the courts to cases involving questionable eyewitness identifications suggests an increased desire to understand the factors that influence reliability in order to prevent wrongful convictions due to misidentifications. Some courts have maintained a focus on system variables to determine admissibility (e.g., Perry v. New Hampshire, 2012), while others have expanded examination of factors to those beyond the control of the police that can influence the likelihood of an accurate identification, especially factors not intuitive to jurors. The *Henderson* Court anticipated that the majority of identifications would still be ruled admissible, yet tasking jurors with determining the reliability of eyewitness identification can be problematic because jurors tend to overestimate the accuracy of eyewitnesses (Devenport et al.,
1997) and are unfamiliar with many factors that influence identification accuracy (Desmarais & Read, 2011). Beyond evaluating eyewitness evidence, some social-cognitive theories may provide an explanation for how jurors reason and evaluate evidence more generally.
Chapter 3: How do Jurors Reason?

Jurors are assigned the arduous task of examining and processing copious amounts of evidence to determine an appropriate verdict. Jurors may have particular difficulty deciphering the accuracy of eyewitness testimony as eyewitnesses may appear very confident in their identification, yet be completely mistaken (Devenport et al., 1997). Furthermore, jurors may be unaware of a variety of factors that can influence eyewitness identification accuracy (Desmarais & Read, 2011). Several social-cognitive theories, including the story model (Pennington & Hastie, 1986), elaboration likelihood model (ELM; Petty & Cacioppo, 1986), and heuristic-systematic model (HLM; Chaiken, 1980) outline the decision making process applicable to jurors and may provide insights into the problems which arise in eyewitness cases.

**Story Model**

In the story model, Pennington and Hastie (1986) propose that jurors seek meaning in the evidence they are given. Jurors will first create stories based on the evidence presented at trial as well as personal knowledge of similar events (Pennington & Hastie, 1986). Jurors' attitudes and beliefs will also influence the way evidence is interpreted (Pennington & Hastie, 1992). Through judicial instructions, jurors are then made aware of the verdict categories available in order to classify one of their stories into a specific verdict category.

Story construction is central to the theory and is what will determine one’s verdict choice (Pennington & Hastie, 1992). If more than one story is constructed, jurors will choose the story that has the best explanation, and is the most consistent, complete, plausible, and unique (Levett, Danielsen, Kovera, & Cutler, 2005). As an illustration of the role of story construction and the influence of personal experience, Huntley and Costanzo (2003) examined four sexual harassment
cases. Story selection mediated the relationship between juror gender and verdicts such that women were more likely to find for the plaintiff than men due to the differing stories that women endorsed.

**Elaboration Likelihood and Heuristic Systematic Models**

The dual-process theories of the elaboration likelihood model (ELM; Petty & Cacioppo, 1986) and heuristic systematic model (HSM; Chaiken, 1980) have been offered to explain how people process persuasive arguments. Systematic processing involves an increased cognitive effort and more detailed evaluation of the content of the message (Chaiken, 1980). In contrast, heuristic processing entails only superficially evaluating information by relying on shortcuts, such as source related cues (Chaiken & Maheswaran, 1994).

In the ELM, the central route to persuasion requires not only more cognitive effort similar to systematic processing, but also the motivation and ability to process the content of a message (Petty & Cacioppo, 1986). If either motivation or ability is lacking, the peripheral route to persuasion is engaged and the content of the message is not evaluated as thoroughly (Petty & Cacioppo, 1986). When this occurs, jurors may use cognitive shortcuts or heuristic cues external to the message content to determine the quality of the argument (Ratneshwar & Chaiken, 1991). For example, within a civil trial context, complex expert testimony on toxic substances may induce jurors to rely on peripheral cues, such as expert pay, to determine the quality of the testimony (Cooper & Neuhaus, 2000). In eyewitness cases, peripheral cues, such as confidence, may be relied upon by jurors during an eyewitness’ testimony despite confidence not being a perfect predictor of identification accuracy (Cutler et al., 1988).
Overall, the theories proposed to explain juror decision making highlight the importance of juror knowledge, motivation, and ability. The story model assumes that if jurors have knowledge they believe is applicable to their decision in a given case, they will use that knowledge to inform their decision making (Pennington & Hastie, 1986). However, juror knowledge concerning eyewitness identification may be incomplete, inaccurate, or both. Additionally, jurors need both motivation and ability to thoroughly evaluate evidence (see ELM; Petty & Cacioppo, 1986). An eyewitness identification may involve a multitude of complex interactive factors that may not be intuitive to jurors. The Henderson court concluded that courts overestimate the ability of jurors to evaluate such factors and additional education is needed to ensure jurors are able to determine the reliability of an identification (New Jersey v. Henderson, 2001). The Henderson court placed emphasis on using educational aides that are based on “reliable scientific evidence that experts generally accept,” which is often achieved through the use of meta-analyses (New Jersey v. Henderson, 2011, p. 882). The following section will review the expansive literature, including meta-analyses, on several eyewitness factors relevant to the current study.
Chapter 4: Eyewitness Identification Research

The New Jersey Supreme Court stated, “We are convinced from the scientific evidence in the record that memory is malleable, and that an array of variables can affect and dilute memory and lead to misidentifications (p. 872).” In formulating their opinion, the *Henderson* Court relied on the growing body of eyewitness identification research published in the past 40 years, recognizing that a variety of factors can influence the accuracy of eyewitness identification. Those factors beyond the control of the criminal justice system are called estimator variables, which concern characteristics of the crime, eyewitness, and perpetrator (Wells, 1978). System variables, on the other hand, concern factors that occur after the crime and usually surround the identification procedure. In the past thirty years, there has been an explosion of research devoted to understanding how these factors influence identification accuracy, most of which were presented to and relied upon by the *Henderson* court in their opinion (*New Jersey v. Henderson*, 2011).

**Estimator Variables**

A variety of variables at the time of a crime can influence an eyewitness' ability to encode, store, and later accurately retrieve key details concerning the perpetrator. The most researched and well-understood estimator variables include alcohol, exposure duration, retention interval, stress, weapon focus, and the other race effect. Alcohol can reduce both the amount and accuracy of information that an eyewitness can recall one week after witnessing a crime (Yuille & Tollestrup, 1990). Furthermore, alcohol increases the likelihood of a false identification in target absent lineups (Yuille & Tollestrup, 1990) and showups (Dysart, Lindsay, MacDonald, & Wicke, 2002).
The length of time a witness had to view the culprit (e.g., exposure duration) and the delay between witnessing the crime and making an identification (e.g., retention interval) influence the quality of an eyewitness' memory for an event (Shapiro & Penrod, 1986). Exposure to a perpetrator for 45 seconds is associated with higher identification rates (90% accurate) compared to only a 12 second exposure (32% accurate; Memon, Hope, & Bull, 2003). Better identification accuracy may be due to the additional time to encode facial information which increases memory strength (Memon et al., 2003). A recent meta-analysis indicated an increase in both correct rejections in target absent lineups and correct identifications in target present lineups with longer versus shorter durations ($r = .30$; Bornstein, Deffenbacher, Penrod, & McGorty, 2012).

Regarding retention interval, research demonstrates that memory for faces deteriorates most rapidly immediately following an event and levels off over time (Deffenbacher, 2008). Specifically, memory strength is reduced by 15% in the first 10 minutes after an event (Deffenbacher, Bornstein, McGorty, & Penrod, 2008). The results of a meta-analysis demonstrate a modest effect of retention interval such that the probability of a correct recognition judgment is decreased as the retention interval increases ($r = .18$; Deffenbacher et al., 2008).

The negative impact of exposure duration and retention interval on eyewitness memory may seem intuitive, but the effect of stress is not as obvious. While counter-intuitive, a face can be forgotten and memory hindered by high levels of stress (Morgan et al., 2004). Specifically, when a perpetrator is present in a lineup, eyewitnesses under high stress are only accurate 39% of the time compared to 59% when an eyewitness is under low stress (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). The results of a meta-analysis found that stress has a moderate effect
on both identification accuracy ($r = -.25$) in target present arrays and modest effect on eyewitness recall ($r = -.15$), decreasing both as stress increases (Deffenbacher et al., 2004).

An eyewitness’ stress level may be influenced by the presence of a weapon during the commission of a crime. Regardless of stress, however, weapon presence can shift attention toward the weapon and away from the perpetrator (Hope & Wright, 2007). This attention shift reduces the ability to encode a perpetrator's face and thus reduces identification accuracy (Steblay, 1992). Supporting this contention, a meta-analysis of 28 studies indicated that weapon presence has a moderate effect on memory performance ($r = -.26$), including feature and identification accuracy such that memory performance was decreased when a weapon was present (Fawcett, Russell, Peace, & Christie, 2013).

The encoding process can also be disrupted when an eyewitness is exposed to a perpetrator of a different race. People generally have difficulty accurately identifying persons of another race relative to their own race (i.e., other race effect; Wells & Olson, 2001). This difficulty may be due to the way faces are processed (Michel, Rossion, Han, Chung, & Caldara, 2006) or categorized (Maclin & Malpass, 2003). For example, Caucasians process other Caucasian faces more holistically than other race (e.g., Asian) faces (Michel et al., 2006). Additionally, categorizing a face as another race can induce stereotypic facial features associated with that race, resulting in a distorted memory (Maclin & Malpass, 2003).

A meta-analysis covering over 30 years of research demonstrated a modest race effect for hits ($r = .11$) and false alarms ($r = .19$; Meissner & Brigham, 2001). The odds of a correct identification of a same race face are 1.4 times more likely than an other-race face, while the odds of a false identification are 1.56 times greater for other race compared to same race faces.
(Meissner & Brigham, 2001). Furthermore, Caucasians (relative to other racial groups) appear particularly inept at identifying and distinguishing between persons belonging to racial minority groups (e.g., African Americans; Meissner & Brigham, 2001).

**System Variables**

Unlike estimator variables, the criminal justice system has control over most system variables. System variables can impact eyewitness memory, but their impact is largely due to the social influence of police procedures, such as lineup administration (Wells et al., 2006). In particular, lineup presentation type, lineup instructions, lineup fairness, confirmatory feedback, and double-blind administration can all influence the suggestiveness of lineups and ultimately the decisions of eyewitnesses (Wells et al., 2006; Steblay, Dysart, Fulero, & Lindsay, 2003; Deffenbacher, Bornstein, & Penrod, 2006).

Police have a variety of ways to present a suspect to an eyewitness, including a simultaneous lineup, sequential lineup, or a showup (Lindsay & Wells, 1985). Police may show photographs or live lineup members all at once (i.e., simultaneously) to an eyewitness. The use of simultaneous lineups induces eyewitnesses to rely on relative judgment (Wells & Olson, 2003). That is, eyewitnesses compare lineup members to each other, choosing the person who most closely resembles the eyewitness’ memory of the culprit. The use of relative judgment is not problematic when the perpetrator is in the lineup as an eyewitness’ memory matches the perpetrator. However, relative judgment increases the likelihood of a false identification when the perpetrator is absent and an innocent suspect who most closely matches the eyewitnesses' memory is chosen instead (Steblay et al., 2011).
Police may also present lineup members to an eyewitness one at a time using a sequential lineup procedure (Lindsay & Wells, 1985). Eyewitnesses shown lineup members one at a time appear to use “absolute judgment” by comparing each member to their memory of the culprit instead of using “relative judgment”, where members of an array are compared to one another (Wells & Olson, 2003). Eyewitnesses are more likely to choose a lineup member in a simultaneous lineup compared to a sequential lineup regardless of whether the perpetrator is present in the lineup (Steblay, Dysart, & Wells, 2011). Thus, relative to a sequential lineup, the use of a simultaneous lineup results in more correct identifications when the perpetrator is present \((r = .08)\), but this small increase in correct identifications may be due to guessing (Steblay et al., 2011). More problematic, when the perpetrator is absent, false identifications are more likely with a simultaneous compared to sequential lineup \((r = .23\); Steblay et al., 2011). These meta-analytic results indicate that identifications from sequential lineups are more diagnostic of guilt and therefore preferable to simultaneous lineups (Steblay et al., 2011).

The showup is a more suggestive type of identification procedure used by police where an eyewitness is aware that only one person will be shown to them (Lindsay & Wells, 1985). In simultaneous and sequential lineups, identification errors may be distributed across known innocents (also known as fillers), but a showup will only contain a suspect who may or may not be innocent (Wells & Quinlivan, 2009). While police know an eyewitness is unreliable when they choose a filler, there is no such protection with a showup procedure (Steblay, Dysart, Fulero, & Lindsay, 2003). In a large meta-analysis, Steblay et al. (2003) reported a small effect size \((r = .07)\) comparing showups to (biased) lineups, where the percentages of errors were 23% and 17%, respectively. However, Steblay et al. (2003) also noted that the differences in
performance would be substantially larger if the lineups were unbiased (15% for showups vs. 7% for lineups).

Prior to an identification procedure, police may give unbiased lineup instructions which explicitly inform the eyewitness that the perpetrator may or may not be present in the lineup (or showup; Malpass & Devine, 1981). In contrast, biased lineup instructions do not explicitly state the possible absence of the perpetrator and may indicate to an eyewitness that the suspect is present and an identification should be made (Steblay, 1997). Unbiased instructions reduce an eyewitness' reliance on relative judgment, whereby an eyewitness chooses the lineup member who most resembles the perpetrator relative to the other lineup members (Wells et al., 1998). The results of a meta-analysis suggest that instructions have a moderate effect on identification accuracy with target-absent lineups ($r = .28$; Steblay, 1997). Specifically, when the perpetrator is absent, biased instructions decrease identification accuracy (35%) compared to unbiased instructions (60%; Steblay, 1997).

To reduce the likelihood of a misidentification, lineup members should be similar in appearance. When an innocent suspect stands out in a lineup, the chances of a false identification are increased (proportion choosing an innocent suspect in low similarity lineups = .70 vs. high similarity = .31; Lindsay & Wells, 1980). Relative to low similarity lineups, high similarity lineups slightly reduce correct identifications but increase diagnosticity (of both guilt and innocence; Lindsay & Wells, 1980). To further increase lineup fairness, lineup members should be matched to the eyewitness’ verbal description of the perpetrator, not to the suspect that will be included in the lineup (Clark & Tunnicliff, 2001). When lineup members are matched to an innocent suspect, the likelihood a false identification increases (probability of choosing an
innocent suspect in suspect matched lineup = .25 vs. perpetrator matched = .05; Clark & Tunnicliff, 2001).

Perhaps the most researched of eyewitness identification factors concerns eyewitness confidence (Wells et al., 2006). Jurors often rely on an eyewitness' confidence as a proxy for the accuracy of an identification (Cutler et al., 1990). Yet, a meta-analysis with over 4,000 participants indicated confidence does not perfectly correlate with identification accuracy ($r = .37$ for choosers and $r = .12$ for non-choosers), thereby making confidence an unreliable indicator of identification accuracy (Sporer, Penrod, Read, & Cutler, 1995).

Confidence also proves to be malleable and susceptible to suggestive police procedures, such as post-identification feedback (Wells & Bradfield, 1998). Confidence becomes inflated, especially for inaccurate eyewitnesses, when a lineup administrator confirms an eyewitness has made the correct choice (e.g., “Good, you identified the actual suspect.”; Bradfield, Wells, & Olson, 2002), but also when more subtle feedback is given from a non-blind administrator (e.g., “Thank you. You have been a really great witness”; Dysart, Lawson, & Rainey, 2012). A meta-analysis of 7,000 participants indicated that confirmatory post-identification feedback not only inflates confidence ($r = .23$ for culprit present lineups; $r = .44$ for culprit absent lineups), but also influences retrospective accounts of an eyewitness' opportunity to view a perpetrator ($r = .24$ for culprit present lineups; $r = .28$ for culprit absent lineups) as well as attention paid to a perpetrator's face ($r = .14$ for culprit present lineups; $r = .23$ for culprit absent lineups; Steblay et al., 2014; see also Douglass & Steblay, 2006), all of which have been identified by the courts as reliable indicators of accuracy (Neil v. Biggers, 1972).
To prevent lineup administrators from conveying to an eyewitness any information relevant to the identity of the suspect, researchers have suggested that police implement double blind administration (Greathouse & Kovera, 2009), whereby lineup administrators are unaware of which lineup member is the suspect (Wells & Luus, 1990). False identifications increase in target absent lineups when the lineup administrator is aware of who the suspect is and nudges the eyewitness toward that lineup member (31.7% when single blind vs. 10% when double blind; Alberts, 2007).

When an eyewitness’ memory for a perpetrator is strong, system variables should have little influence (Wells et al., 2006). However, when various estimator variables (e.g., poor lighting, disguise) hinder memory, the impact of system variables will likely be larger (Wells et al., 2006). That is, suggestive police procedures, such as confirming an eyewitness’ lineup selection or failing to inform an eyewitness that a suspect may or may not be present in a lineup, can alter an eyewitness’ weak memory for an event as well as confidence in one’s memory, and thus increase the probability of a mistaken identification.

Estimator and system variables rarely operate in isolation. Furthermore, the presence of some of these variables may exacerbate the problematic nature of others. For example, cross-race identification may be moderated by both exposure time and the time delay between witnessing a crime and identification (Meissner & Brigham, 2001). Shorter exposure times increase the false identification rate of other race faces (Meissner & Brigham, 2001). And as time increases, forgetting is more rapid for cross-race faces than same-race faces which are less affected by the interference of other faces (Deffenbacher et al., 2008). Additionally, the behavior of single blind lineup administrators is more influential when biased lineup instructions and a simultaneous
lineup are also used (Greathouse & Kovera, 2009). That is, single blind lineup administrators increase the likelihood that an eyewitness will choose a suspect (who may or may not be innocent) when biased lineup instructions are given for a simultaneous lineup (Greathouse & Kovera, 2009). Taken together, many factors can contribute to a mistaken identification. Should an identification be admitted in court, it becomes the jury’s responsibility to determine whether such factors impaired the ability of an eyewitness to make an accurate identification.
Chapter 5: Juror Knowledge Regarding Eyewitness Identification Factors

Much research has addressed the factors that influence the likelihood of an accurate identification. Unfortunately, jurors’ knowledge of these factors is limited. A recent meta-analysis assessing lay knowledge indicated that 75% of 16 factors known to influence eyewitness identification accuracy are not common sense to jurors (Desmarais & Read, 2011; see Appendix B). In addition, jurors and experts differ in their knowledge of 87% of 30 eyewitness factors (Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006). Some factors, such as weapon focus, cross-race bias, lineup presentation format, and accuracy-confidence relationship appear particularly difficult for the majority of jurors to understand, while other factors are less problematic (e.g., alcohol intoxication and confidence malleability; Desmarais & Read, 2011).

Knowledge Integration

Beyond acquiring knowledge of factors that influence identification accuracy, jurors must also be able to integrate such knowledge into their decision making. A few studies have found jurors capable of such a task. For example, one study found that those who have an understanding of the weapon focus effect are able to correctly reduce their perception of an eyewitness’ credibility when a weapon is present compared to absent (Neal, Christiansen, Bornstein, & Robicheaux, 2012). Another study established that mock jurors (both undergraduates and community members) understand and can apply knowledge of foil biased lineups (i.e., perpetrator description matches suspect, while other lineup members are poorer matches; Devenport, Stinson, Cutler, & Kravitz, 2002). Jurors rated foil biased lineups as more suggestive and less fair and the defendant as less culpable than when foil unbiased lineups are used.
Others have demonstrated that jurors are sensitive to biased lineup instructions. Cutler and colleagues (1988) presented undergraduates with a trial involving an eyewitness identification where nine estimator (e.g., disguise) and system variables (e.g., lineup instructions) in addition to witness confidence were manipulated. Guilty verdicts were higher when jurors were informed that the police used unbiased compared to biased lineup instructions. However, eyewitness confidence also influenced verdicts regardless of the quality of other system and estimator variables (Cutler et al., 1988). When the eyewitness was 100% confident, both guilty verdicts and the probability that the identification was correct increased compared to when the eyewitness was only 80% confident.

Replicating the Cutler et al. (1988) study using jury eligible community members, Cutler and colleagues (1990) found that confidence was the only factor that affected jurors’ verdicts and probability of an accurate identification. The influential nature of eyewitness confidence is problematic considering the aforementioned literature has demonstrated that confidence does not always indicate accuracy (see Sporer et al., 1995). These results suggest jurors are insensitive to most factors known to influence eyewitness identification accuracy and are instead relying on a heuristic cue (i.e., eyewitness confidence) to determine the quality of the eyewitness testimony.

Jurors also tend to overestimate the ability of others to make accurate identifications as well as have difficulty discriminating between accurate and inaccurate eyewitnesses themselves (Devenport et al., 1997). In one study, jury-eligible community members were asked to review two eyewitness identification research studies and then estimate how many participants made a correct identification from a target present lineup (Brigham & Bothwell, 1983). The vast majority (i.e., 70-90%) overestimated the correct identification rate for each study. Additionally,
when eyewitnesses were very confident, jurors ignored witnessing conditions when determining the accuracy of the identification (Lindsay, Wells, & Rumpel, 1981). When an eyewitness’ confidence was low and jurors were informed of unreliable witnessing conditions (i.e., short exposure and a disguised perpetrator), belief in the accuracy of that eyewitness was reduced (Lindsay et al., 1981).

In short, lay knowledge concerning factors that influence eyewitness identification accuracy is inadequate (Desmarais & Read, 2011). When such knowledge is acquired, it is rarely applied to evaluate specific case facts when rendering a verdict (Cutler et al., 1988). Instead, jurors appear to rely on eyewitness confidence as an indicator of the quality of an identification (Lindsay et al., 1981; see also Penrod & Cutler, 1995). As a result, additional education may be needed to increase juror sensitivity to factors that influence eyewitness identification accuracy.
Chapter 6: The Efficacy of Traditional Safeguards

The courts have established a number of safeguards to protect against jurors relying on unreliable evidence such as mistaken eyewitness identifications. These include voir dire, cross-examination, expert testimony, and judicial instructions. The effectiveness of each safeguard is dependent upon maximizing juror sensitivity to factors that influence the accuracy of an identification without inducing a general sense of skepticism regarding the ability of all eyewitnesses to make correct identifications (Cutler, Dexter, & Penrod, 1989). In other words, safeguards that sensitize jurors would result in increased convictions when eyewitness conditions are good and decreased convictions when eyewitness conditions are poor (Cutler et al., 1989). In contrast, skepticism occurs when jurors reduce convictions regardless of the quality of an eyewitness identification.

Voir Dire

Traditionally, voir dire has been used to identify and eliminate from the jury pool people with biases that could influence verdict decisions. During the process of voir dire, attorneys may request that a juror be struck due to some demonstrable bias (i.e., striking for cause) or for no specified reason (i.e., through peremptory challenges but not for unconstitutional reasons such as race). Attorneys should seek to identify potential jurors who are not willing or able to thoroughly evaluate eyewitness evidence (Narby & Cutler, 1994). Unfortunately, attitudes concerning trust in eyewitness testimony are poor indicators of verdict, suggesting that voir dire may not be a particularly effective safeguard (Narby & Cutler, 1994).

Cross-examination
Cross-examination is a commonly used safeguard in adversarial settings to identify weaknesses in the opposing side's case. Attorneys must first understand the factors that affect eyewitness identification accuracy in order to ask appropriate questions that highlight the quality of the identification (Devenport et al., 1997). Defense attorneys appear better educated about eyewitness identification factors compared to prosecutors, but knowledge on factors such as the forgetting curve and lineup presentation format could still be improved (Wise, Pawlenko, Safer, & Meyer, 2009).

Even if attorneys understand and can argue the factors influencing identification accuracy in a particular case, jurors must also understand and integrate such knowledge into their decision making. One study examined the effectiveness of cross-examination by having participants view taped direct and cross-examinations of eyewitnesses who observed a staged crime and later made known correct or incorrect identifications (Lindsay, Wells, & O'Connor, 1989). The direct and cross-examinations were conducted by either experienced attorneys or senior law students. Verdict was unrelated to eyewitness accuracy and experienced attorneys were no more likely to assist mock jurors in distinguishing between accurate and inaccurate eyewitnesses than the senior law students. The failure of cross-examination as an effective safeguard could be due to its adversarial nature that seeks to draw out any deception from witnesses, but fails to detect when a witness is honest, but mistaken (Epstein, 2007).

**Expert Testimony**

Another potential safeguard against erroneous convictions is expert testimony. Expert testimony may be admitted at trial if it "will assist the trier of fact to understand the evidence or to determine a fact in issue” (Federal Rules of Evidence (FRE), 1975, p. 14). Since 75% of
factors known to influence eyewitness identification accuracy are not common sense to jurors, expert testimony on these factors could assist jurors in their decision-making (Desmarais & Read, 2011). Indeed, experts agree that many of these eyewitness factors (e.g., the other race effect, weapon focus, confirmatory feedback) are worthy of testifying about in court (Kassin, Tubb, Hosch, & Memon, 2001).

The goal of expert testimony is twofold (Leippe & Eisenstadt, 2009). The first goal is to reduce overreliance or over-acceptance of eyewitness identifications. The second goal is to educate jurors on the factors that influence eyewitness memory and accuracy (Leippe & Eisenstadt, 2009). Research on the effectiveness of expert testimony has been mixed (see Martire & Kemp, 2011), identifying sensitivity effects, skepticism effects, or null effects. Expert testimony may sensitize jurors to some eyewitness identification factors, such as biased lineup instructions (Cutler et al., 1989; Devenport et al., 2002), weapon focus, and witness confidence (Cutler et al., 1989). When an expert testified about the influence of these factors on eyewitness accuracy, jurors adjusted their perception of the suggestiveness of the lineup procedures used to obtain the identification (Devenport et al., 2002) as well as the accuracy of the identification and verdict (Cutler et al., 1989) according to the quality of the witnessing and identification conditions.

Other studies have found a skepticism effect with the use of expert testimony. For example, jurors exposed to the testimony of a court-appointed expert rated the prosecution’s case as weaker, were less likely to believe that the eyewitness correctly identified the perpetrator, and were less likely to convict compared to those who did not hear expert testimony (Cutler et al., 1990). When a court-appointed expert testified, skepticism of the eyewitness evidence remained
regardless of witnessing and identification conditions (e.g., use of a disguise, weapon). Jurors in this study may have relied on the court-appointed nature of the expert as a heuristic cue to credibility instead of thoroughly evaluating the testimony (Cutler et al., 1990; see also Chaiken, 1980). Yet, some evidence of skepticism has been replicated with the use of an adversarial expert. Specifically, undergraduates who viewed defense expert testimony on eyewitness identification were less likely to believe that the defendant was guilty and the eyewitness' identification was accurate compared to those who did not view expert testimony (Fox & Walters, 1986). In that study, however, witness and identification conditions were not manipulated leaving open the possibility that the observed skepticism effects may have actually reflected sensitivity to poor eyewitness evidence.

Still, others have found null effects. In one study, both undergraduates and community members were exposed to a defense expert, opposing experts, or no expert testimony concerning instruction and foil biases during a lineup identification procedure (Devenport & Cutler, 2004). Neither a defense-only expert nor opposing experts sensitized jurors to variations in instruction and foil biases or made jurors skeptical of the eyewitness regardless of the quality of the identification (Devenport & Cutler, 2004). Additionally, Martire and Kemp (2009) manipulated the presence of expert testimony concerning an eyewitness identification where ground truth was established (i.e., the researchers knew whether or not the eyewitness was accurate). Expert testimony did not affect jurors’ ability to distinguish accurate from inaccurate eyewitnesses. Verdict differences were not reported.

The research is not purely favorable concerning the overall effectiveness of expert testimony. In fact, a review of the research indicates that skepticism is the most common effect
observed in studies of expert testimony (Martire & Kemp, 2011). Yet, there appears to be more support for this safeguard than other, more traditional safeguards (e.g., voir dire, cross-examination; Desmarais & Read, 2011). Any sensitivity effects produced by expert testimony may be due to the unique way (relative to other safeguards) experts educate jurors who lack knowledge about eyewitness issues (Desmarais & Read, 2011).

The specificity of an expert’s testimony may also influence juror decisions. For example, more general expert testimony on memory processes reduces the perceived accuracy of an eyewitness’ identification compared to those who do not hear expert testimony at all, but to a lesser extent than case-specific testimony (Fox & Walters, 1986). In contrast, another study found that sensitivity to witnessing and identification conditions decreased with more detailed compared to more general expert testimony (Cutler, Penrod, & Dexter, 1989b). General expert testimony consisted of a description of relevant factors that influence identification accuracy, while the detailed, research-enhanced expert testimony included the general testimony plus an explanation of the changes in rates of correct and false identifications as a function of these factors. Cutler and colleagues surmised that jurors might assume the relationships between various eyewitness factors and accuracy are larger than they actually are when evaluating this evidence on their own. Additional research is needed to determine how this specific type of testimony, presented via judicial instructions, might impact jurors’ decisions.

**Judicial Instructions**

Relative to other safeguards, legal scholars tend to prefer judicial instructions as they are less costly and time consuming compared to experts (Sheehan, 2011). Presented with information outlined in the previous sections, the *Henderson* Court proposed judicial instructions
as one way to educate jurors about factors that influence eyewitness identification accuracy. To be effective, jurors must be willing and able to follow the instructions (Devine et al., 2001). Yet, much research attests to the difficulty jurors have in understanding and applying a wide variety of judicial instructions, such as pattern instructions (Reifman et al., 1992), those to disregard or limit evidence (Steblay et al., 2006), and instructions on eyewitness identification (Greene, 1988; see Bornstein & Hamm, 2012 for review).

Jurors typically receive instructions on both substantive (i.e., the specific elements of a crime) and procedural (e.g., reasonable doubt, presumption of innocence in criminal trials) rules of law. While understanding of procedural rules is higher among instructed versus uninstructed jurors, understanding of substantive rules is not improved (Reifman et al., 1992). Additionally, overall comprehension of such instructions is rarely above chance levels (Reifman et al., 1992; see also Lieberman & Sales, 1997), even with the aid of deliberation (Ellsworth, 1989).

To improve jurors’ comprehension of instructions, a few techniques have been employed. For example, while still low, comprehension rates do appear to increase when instructions are simplified, including death penalty instructions (Frank & Applegate, 1998), pattern instructions (Severeance & Loftus, 1982), and eyewitness instructions (Greene, 1988). Additionally, providing written copies of death penalty instructions to jurors increases comprehension, though only when those instructions are simplified (Frank & Applegate, 1998). Others have not observed improved comprehension with written pattern instructions, though participants in a field study reported being more satisfied with written instructions (Heuer & Penrod, 1989). Finally, the timing of instructions may influence jurors’ evaluation of the evidence. Instructions are typically administered at the conclusion of the trial after all evidence has been presented.
Providing instructions on reasonable doubt and burden of proof at the outset of the trial, however, can increase jurors’ threshold for conviction (Kassin & Wrightsman, 1979) and assist jurors in evaluating evidence in line with legal standards (Heuer & Penrod, 1989).

The ability of judges to explain or educate the jury will influence the effectiveness of various instructions, including eyewitness instructions. Such instructions should sensitize jurors to factors that influence the accuracy of eyewitness identification. The *Telfaire* instruction is the most commonly used instruction cautioning jurors about the problems with eyewitness identification (*US v. Telfaire*, 1972). This instruction is inadequate for several reasons. First, jurors are not provided with an explanation as to how memory works nor how the five factors outlined in the instruction influence eyewitness accuracy. Second, the five factors are not based on psychological research, but on previous case law (Cutler & Penrod, 1995). The factors referenced are the opportunity of the witness to view the criminal at the time of the crime, the witness’ degree of attention, the accuracy of his prior description of the criminal, the witness’ certainty, and the time between the crime and the identification (*Neil v. Biggers*, 1972). Most problematic of these factors is eyewitness certainty. The relevant literature has demonstrated the complex relationship between accuracy and eyewitness confidence, which is moderated by the presence of other factors not outlined in the *Telfaire* instruction (e.g., confirmatory feedback; see Wells et al., 2006). Finally, the absence of other factors known to influence eyewitness accuracy makes it difficult for jurors to draw accurate conclusions about the reliability of an identification.

The *Telfaire* instruction fails to increase comprehension concerning the factors contained in the instruction. One study found that jurors given the *Telfaire* instruction were no more likely
to understand the problems in an eyewitness identification than jurors who do not hear the instruction (Greene, 1988). Additionally, the believability of the eyewitness did not vary depending on the presence of the Telfaire instruction or the quality of the witnessing and identification conditions. Others have found similar null effects such that the Telfaire instruction had no appreciable effect on verdict, the perceived accuracy of the identification, or the eyewitness' credibility (Cutler et al., 1990). Similarly, using a more ecologically valid design, Martire and Kemp (2009) found that eyewitness instructions do not influence juror decisions. A recent meta-analysis supports the contention that eyewitness judicial instructions do not sensitize jurors to the quality of eyewitness identification (Berman & Penrod, 2015).

Greene (1988) attempted to revise the Telfaire instruction to simplify the language, clarify the relationship between confidence and accuracy, and elaborate on additional factors not included in the original instruction. Comprehension was improved with the use of the revised instruction relative to the original instruction and no instruction. Instead of the desirable effect of sensitivity, however, the revised instruction decreased convictions regardless of the quality of the witnessing and identification conditions. Thus, the revised instruction made jurors skeptical of all eyewitness identifications.

Others have also found skepticism effects. For example, Laub, Kimbrough, and Bornstein (2014) found that, regardless of witnessing and identification conditions, those receiving abbreviated, research-supported witness instructions reduced perceptions of witness competence, but not verdict. Ratings of witness confidence were also reduced with jury instructions, though only when the quality of the identification was poor. The instructions used in this study focused on disguise, weapon focus, exposure duration, retention interval, and witness confidence (not all
of which were manipulated), suggesting that general witness instructions may not be sufficient to sensitize jurors to eyewitness evidence quality. Even case-specific eyewitness instructions may induce skepticism. Participants in one study were given eyewitness instructions via a simple fact sheet concerning factors relevant to a case that research has demonstrated to influence eyewitness accuracy (e.g., weapon focus; Moore, 2011). Mock jurors reduced conviction rates when given this fact sheet at the end of a trial vignette, but not the beginning (Moore, 2011). Methodologically, as with Fox and Walters (1986), evidence was not manipulated. Therefore, whether the effects were due to sensitivity to poor eyewitness evidence or to skepticism cannot be determined. Unlike Moore (2011), initial evaluations of the case-specific Henderson instructions suggests it creates skepticism regardless of when the instruction is given. Henderson instructions administered either prior to an eyewitness’ testimony or at the end of trial induced skepticism of the eyewitness’ accuracy, reducing conviction rates regardless of the quality of the witnessing and identification conditions (Berman et al, 2015). Similar results (i.e., skepticism) were found comparing the Henderson instructions to standard eyewitness instructions used in Florida (Yokum & Papailiou, 2014).

More recently, Pawlenko, Safer, Wise, and Holfeld (2013) tested the effectiveness of the interview-identification-eyewitness (I-I-Eye) teaching aid administered at the beginning of a trial before any evidence was introduced. This aid relies on established psychological research findings and focuses on how police conduct interviews with eyewitnesses as well as the quality of estimator and system variables not specific to a case. The aid is very brief, repetitive, and rhetorically asks jurors to evaluate the quality of various identification factors (e.g., For how long did the eyewitness view the perpetrator?). Compared to participants receiving general
information about the trial process or the Biggers criteria, those who received the I-I-Eye teaching aid demonstrated increased sensitivity to system variable manipulations (e.g., confirmatory feedback; Pawlenko et al., 2013). Only with the aid of the I-I-Eye did participants increase conviction rates when the evidence was strong and decrease conviction rates when the evidence was weak.

Across studies of different safeguards, researchers have assessed a variety of outcomes, including how much the eyewitness is believed and verdict. Leippe and Eisenstadt (2009) suggest there are both proximate (educational) and distal variables of interest. They point out that researchers have often failed to examine the educational function of these safeguards (e.g., increased comprehension, adjusted beliefs in eyewitness accuracy), which they suggest may be intermediary variables that facilitate changes in verdict. The current studies address both proximate and distal variables of interest.

Overall, the majority of proposed safeguards do not serve to protect against wrongful convictions due to unreliable eyewitness evidence. Specifically, voir dire and cross-examination appear ineffective at sensitizing jurors to eyewitness identification problems. While, expert testimony appears to sensitize jurors to a few factors, such as weapon focus (Cutler et al., 1989), it is not effective for other factors (e.g., biased lineups; Devenport & Cutler, 2004). Furthermore, most studies on eyewitness instructions, including case-specific Henderson instructions, suggest instructions lead to skepticism. The one eyewitness instruction study finding sensitivity effects was not case-specific and did not manipulate the presence of estimator variables so it is unknown whether similar instructions could demonstrate sensitivity to these factors (Pawlenko et al., 2013). Because initial findings suggest the Henderson instructions are rather ineffective in their
current state (Berman et al., 2015), a “research enhanced” version may increase juror sensitivity.

Sensitivity effects observed in cases with expert testimony may be due to the specific research (typically meta-analyses) discussed by experts during their testimony. Some differences in the effectiveness of expert testimony appear to exist depending on an expert’s specificity (e.g., Cutler et al., 1989), but it is unknown whether the addition of specific research findings to the *Henderson* instructions would better enable jurors to evaluate the evidence relative to instructions without this additional information. Finally, few studies have directly compared the effectiveness of instructions to expert testimony (cf., Martire & Kemp, 2009). The first study seeks to address these shortcomings.
Chapter 7: Conclusion

Various social-cognitive theories have been used to explain juror decision making. Jurors may develop one or more stories to explain the evidence, taking into consideration their own knowledge and attitudes (Story Model; Pennington & Hastie, 1986). Evidence that is complex or not well understood may result in jurors processing information in a simplistic fashion by relying on shortcuts to determine the quality of the evidence (ELM; Petty & Cacioppo, 1986; HLM; Chaiken, 1980).

Eyewitness identification may be one form of complex evidence. A variety of factors influence eyewitness identification accuracy, generally categorized as estimator or system variables (Wells, 1978). Estimator variables, such as exposure duration, stress, the other-race effect, and weapon focus influence an eyewitness' memory and ability to accurately recall crime details (Wells & Olson, 2003). Additionally, police can create suggestive circumstances through the use of system variables, such as a showup or biased lineup instructions, that imply a suspect is present and should be identified (Steblay et al., 2003; Steblay, 1997). Police can also influence an eyewitness' memory and confidence by giving feedback after an identification is made (Dysart et al., 2012).

The New Jersey Supreme Court recognized that many factors that influence eyewitness accuracy are not intuitive to jurors (e.g., weapon focus; see also Desmarais & Read, 2011). Even when jurors have such knowledge, it may not be applied to one's verdict decision (see Cutler et al., 1988). To assist jurors, the courts suggest a variety of safeguards, including voir dire, cross-examination, expert testimony, and judicial instructions. The goal of these safeguards is to sensitize jurors to variations in the quality of an eyewitness identification. Expert testimony
appears to provide jurors with sensitivity to some factors, such as biased lineup instructions (Devenport et al., 2002) and lineup fairness (Cutler et al., 1989). The remaining safeguards have proven ineffective thus far.

The Telfaire instructions, intended to assist jurors with eyewitness evidence, have not affected jurors’ comprehension of such evidence or their verdicts (Greene, 1988; Cutler et al., 1990). The Telfaire instructions may have been ineffective, in part, because the criteria involve estimator variables, which appear especially difficult for jurors to understand (Desmarais & Read, 2011). Furthermore, one of the criteria discussed as a reliable indicator of accuracy is confidence. Jurors may rely on an eyewitness’ confidence as a heuristic cue to accuracy, despite the less than perfect correlation between confidence and identification accuracy (Cutler et al., 1988). Other, more complex factors, such as stress and the other race effect are not discussed or explained by these instructions.

Unlike the Telfaire instructions, which contain a standard set of factors, the Henderson instructions admonish jurors to thoroughly evaluate only the factors considered problematic in a specific case; factors that are based on sound psychological research overwhelmingly supported by the scientific community. Eyewitness instructions must go beyond simply defining an eyewitness factor to explaining how that factor influences identification accuracy. Providing relevant scientific findings may also assist jurors in evaluating eyewitness evidence; this is an empirical question that will be examined in this study.
Chapter 8: Instructions and Expert Testimony

The first study examined the effectiveness of the *Henderson* instructions (see Appendix A). As courts continue to issue new eyewitness instructions, it is necessary to determine whether such instructions induce sensitivity, skepticism, or simply have no impact on jurors’ decisions. Since initial findings indicate the *Henderson* instructions induce skepticism, one potential way to improve their effectiveness would be to elaborate on research speaking to the impact of system and estimator variables on eyewitness accuracy. Such research (typically meta-analyses) is often relied upon by experts and might explain sensitivity effects observed in cases with expert testimony. However, it is unknown whether the addition of specific research findings to the *Henderson* instructions may better enable jurors to evaluate the evidence relative to instructions without this additional information. Providing research findings may increase jurors’ understanding of the relative impact of each factor on identification accuracy, though the results of Cutler and colleagues (1989b) suggest such elaboration could actually reduce sensitivity. Additionally, understanding the relative effectiveness of the *Henderson* instructions compared to expert testimony will assist courts in deciding whether instructions are a sufficient replacement for expert testimony as the *Henderson* Court suggests. The results from this study will likely inform future research that seeks out the most effective and efficient ways to educate jurors on the complex facets of eyewitness identification, especially those that are most counter-intuitive.

To accomplish this, the first study had mock jurors view a videotaped criminal trial that included an eyewitness. The eyewitness witnessed a crime and made an identification under good or poor witnessing and identification conditions. Additionally, jurors heard *Henderson*
instructions, expert testimony, both Henderson instructions and expert testimony, research based Henderson instructions, or no safeguards.

Since recent findings suggested the Henderson instructions are ineffective in their current form, I expected that the research-based instructions which convey information about the magnitude of the effect of factors on eyewitness performance may be on par with expert testimony as the New Jersey Supreme Court envisioned. Therefore, I anticipated that the research-based Henderson instructions and expert testimony would be equally effective at improving jurors’ comprehension for eyewitness factors and sensitizing jurors to variations in witnessing and identification quality. I expected an interaction between quality of witnessing and identification conditions, and safeguard type such that guilty verdicts would be lowest when witnessing and identification quality was poor and highest when witnessing and identification quality was good only when jurors heard the research-based Henderson instructions or expert testimony. I anticipated that these effects would be mediated by witness ratings and belief in the eyewitness. I also explored the impact of having both Henderson instructions and expert testimony on jurors’ decisions. Finally, I compared community and college samples to examine any potential differences in these hypothesized relationships, although past studies suggest this is unlikely (Bornstein, 1999).
Chapter 9: Pilot Study

Pilot testing was initially conducted to examine the stimulus trial which served as the vehicle for testing instruction and expert testimony effects. First, the Innocence Project provided both videotapes and transcripts of an actual trial where the defendant was wrongfully convicted based, in part, on faulty eyewitness identification (Kentucky v. Gregory, 1993). I condensed the trial materials down to key facts surrounding the eyewitness identification, serologist, and alibi witness, which took about three months to accomplish. Next, I hired actors to play the roles of judge, attorneys, and witnesses and filmed over the course of two weekends. Videotaped variations in witnesses, testimony content, and opening and closing arguments were edited to allow for manipulation of as many as eight eyewitness factors and the content of safeguards. The editing process took approximately three months to complete.

The conditions designed to lead to the lowest and highest convictions rates were piloted. Initially, system and estimator variables were manipulated individually, but were eventually grouped to increase the difference between conditions. In the lowest conviction condition, the crime and identification occurred under a variety of bad conditions and neither Henderson instructions nor expert testimony were included. In the highest conviction condition, the crime and identification occurred under a variety of good conditions and both Henderson instructions and expert testimony were included.

A total of 375 (278 undergraduates, 30 community members, and 67 mTurkers) participants were recruited during this phase. Extensive pretesting of the trial was conducted to ensure the overall weight of the evidence that was constant did not clearly support the defense or prosecution. After eight rounds of pilot testing over four months, I acquired an adequate verdict.
split. In the high conviction condition where various estimator and system variables were good, 50% voted guilty. In the low conviction condition where those same variables were poor and both instructions and expert testimony were given, 25% convicted. These conviction rates were lower than anticipated despite the fact that the defendant in the actual case was convicted based on similar evidence. However, I wanted to remain as faithful as possible to the original case.

While I confounded the presence of the *Henderson* instructions and expert testimony, the purpose of piloting was to ensure that verdicts were not at the ceiling or floor and that there was sufficient variation in verdicts to detect differences, if they exist, between conditions. The pretesting phase helped to develop the 20 conditions contained in the first study. The entire stimulus development and piloting phase took approximately one year to complete.
Participants

A total of 209 community members and 243 undergraduate students participated in the study. Data collection from two populations allowed for a direct comparison between community members and undergraduates to assess potential differences in knowledge or knowledge integration. Participants were pre-screened to ensure jury eligibility (i.e., 18 years of age or older and a US citizen). Community members were paid $20, while students received course credit for their participation. Overall, 61% of participants were female and averaged 27.38 years old ($SD = 12.47$). This was a diverse sample (30% White, 29% Hispanic, 23% African American, 17% Other). Ethnicity did not vary by sample type. However, community members were significantly older ($M= 35.71, SD = 13.85$) than students ($M= 20.23, SD = 3.69$), $t(233) = -15.68$, $p < .001$ and included significantly fewer females (52%) than students (69%), $t(428) = -3.83$, $p < .001$.

Design and Materials

This study utilized a 2 (Quality of witnessing conditions: good vs. poor) X 2 (Quality of identification conditions: good vs. poor) X 5 (Safeguard type: none vs. Henderson instructions only vs. research--based instructions only vs. expert testimony only vs. Henderson instructions + expert testimony) fully-crossed factorial design.

Voir dire. Prior to viewing the trial film, participants completed a short voir dire questionnaire. These items included gender, age, citizenship, voter registration, ethnicity, political views, and jury duty history.

Trial stimulus. Mock jurors viewed a criminal trial simulation that ranged from 40-75 minutes in which the prosecution alleged that the defendant committed an attempted rape. This
trial was based on the overturned case, *Kentucky v. Gregory* (1993). The Innocence Project worked to secure the defendant’s exoneration and provided DVDs of the original trial to aide in developing the materials used for the current study.

Filming took place in a mock courtroom using professional actors. All participants viewed New York standard judicial instructions, opening and closing statements from both attorneys, and direct and cross examination of the detective, eyewitness, serologist, defendant, and alibi witness. The presence of *Henderson* eyewitness instructions, research-based *Henderson* eyewitness instructions, and expert testimony varied. Cross-race identification, stress, and low lighting were constant estimator variables across conditions.

**Independent Variables**

**Quality of witnessing conditions.** A block of three estimator variables was manipulated: exposure duration, weapon presence, and time delay. When witnessing conditions were good, the eyewitness saw the perpetrator’s face for 45 seconds, no weapon was present, and there was only one day between when the crime occurred and when an identification was made. When witnessing conditions were poor, the eyewitness only saw the perpetrator’s face for 10-12 seconds, a weapon was present, and there was a one month delay between when the crime occurred and when an identification was made.

**Quality of identification conditions.** A block of three system variables was manipulated: identification procedure type, lineup instructions, and confirmatory feedback. When identification conditions were good, the eyewitness made an identification from a six person lineup where the lineup administrator gave unbiased lineup instructions and did not give the eyewitness any confirmatory feedback after she made an identification. When identification
conditions were poor, the eyewitness made an identification from a showup where the lineup administrator failed to explicitly instruct the eyewitness that this may or may not be the perpetrator and confirmed she identified the suspect before obtaining the eyewitness’ confidence level.

**Safeguard types.** Four safeguards were manipulated in addition to a control group (no safeguards). Participants in the *Henderson* instructions condition heard *Henderson* instructions *only* on the manipulated estimator and system factors that were problematic as the *Henderson* Court desired these instructions to be tailored to the facts of the case. For example, when witnessing conditions were poor, participants heard instructions on duration, weapon presence, and time delay. When identification conditions were poor, participants heard instructions on showups, lineup instructions, and confirmatory feedback. In all conditions, participants heard instructions on double blind administration, confidence and accuracy, cross-race identification, stress, and lighting which were constant factors across all trials (see Appendix A for full instructions).

Participants in the research-based instructions condition also heard *Henderson* eyewitness instructions *only* on problematic estimator and system factors present in the trial they viewed. Additionally, the judge elaborated on research findings specifically discussing how each factor can influence the likelihood of an accurate identification (see brackets in Appendix A for research enhancements).

Like the aforementioned instructions, expert testimony was also case specific. The research discussed in both the research-based instructions and expert testimony remained constant. Lastly, participants in the *Henderson* instructions + expert testimony condition heard
both expert testimony and instructions (without the research-enhancement to avoid being too repetitive) on the same poor estimator and system variables. Participants heard eyewitness instructions prior to the eyewitness’ testimony.

**Mediators**

**Witness ratings.** Participants rated their impressions of the detective and eyewitness on a series of seven-point bipolar adjective pairs. These items include trustworthy-untrustworthy, honest-dishonest, not believable-believable, convincing-unconvincing, and certain-uncertain. For each scale, I recoded items as necessary so that higher values represented higher credibility and averaged the items to create a single scale measure (detective, $\alpha = .87$; eyewitness, $\alpha = .84$).

**Belief in eyewitness.** Participants indicated the probability that the eyewitness correctly identified the defendant as the perpetrator (0% = not at all probable, 100% = completely probable).

**Dependent Measures**

**Verdict.** Mock jurors provided dichotomous judgments of guilt (0 = not guilty; 1 = guilty).

**Comprehension ratings.** Mock jurors indicated their perceptions of whether each of a total of 5 eyewitness factors generally influence the likelihood of an accurate identification using nine-point Likert scales (1 = Strongly Reduces Accuracy, 9 = Strongly Increases Accuracy). These factors included exposure duration, weapon presence, time delay, identification procedure type, and lineup instructions. For example, “As compared to an eyewitness who has 10 seconds to view a perpetrator’s face, what impact would having 45 seconds to view a perpetrator’s face have on the accuracy of an identification?” Jurors also indicated whether confirmatory feedback
generally increases an eyewitness’ confidence. Items were recoded as necessary so that higher values were indicative of increased comprehension.

**Instructions help.** Mock jurors rated the extent to which they agreed the judge’s instructions helped them to evaluate the eyewitness testimony using a nine-point Likert scale (1 = Strongly Disagree, 9 = Strongly Agree).

**Manipulation Checks**

Participants completed 10 manipulation checks by indicating the type of identification procedure used by police, the time delay between witnessing the crime and making an identification, and the amount of time the victim had to see the perpetrator’s face. Participants also indicated whether a weapon was present, and whether the police used biased lineup instructions or gave confirmatory feedback. Lastly, participants answered whether the judge gave instructions on system and estimator variables as well as whether an eyewitness expert gave testimony on system and estimator variables.

**Procedure**

Participants came to the psychology lab to complete the study. After providing informed consent, they completed the voir dire questionnaire. I grouped participants in randomly assigned sessions to view the filmed trial. In the relevant conditions and prior to the eyewitness’ testimony, the video was paused and the experimenter handed out printed copies of the *Henderson* (or research-based) instructions and informed participants that they could follow along with the judge’s instructions. The video resumed with participants hearing the *Henderson* (or research-based) instructions. Participants then individually completed post-trial questionnaires where they determined verdict, belief in the eyewitness, witness ratings,
comprehension ratings, and finally manipulation checks. Participants were subsequently debriefed and excused. Sessions were, on average, two hours.

**Data Analytic Plan**

I first conducted a $\chi^2$ analysis to examine how participants performed on the manipulation checks. Next, I conducted a series of ANOVAs to examine whether participant type differentially influenced witness ratings, comprehension ratings, belief in eyewitness, how helpful instructions were rated, and verdict. Next, I conducted one-way ANOVAs to examine the main effect of safeguard type on comprehension ratings. I then ran ANOVAs to determine if safeguard type moderated the relationship between the evidence manipulations and outcomes (i.e., witness ratings, belief in eyewitness, instructions help, and verdict). Finally, I conducted mediation analyses (PROCESS Model 4, see Hayes, 2013) to examine whether the proposed mediators could explain the effects of the manipulated variables on verdict. With the PROCESS models, I report unstandardized regression coefficients as recommended by Hayes (2013).
Chapter 11: Study 1 Results

Manipulation Checks

Manipulation checks on identification procedure type, lineup instructions, and confirmatory feedback (i.e., quality of identification conditions) were successful with correct responses ranging from 77% to 92%. Participants were also successful at correctly distinguishing poor from good witnessing conditions (i.e., exposure duration, time delay, and weapon presence), with correct responses ranging from 88%–99%. Regarding the safeguard type manipulation, participants were successful at recognizing when the judge gave instructions on both system and estimator variables, with correct responses ranging from 81%–89%. Participants also correctly noted whether or not an expert testified about system and estimator variables, with correct responses ranging from 77% to 89%. Analyses were conducted with the full sample and compared to analyses without participants who failed more than half of the 10 manipulation checks (n=16). None of the results differed; therefore the full sample was retained.

Participant Type

I ran ANOVAs to test for the main and interactive effects of participant type with my other manipulated variables on witness ratings, comprehension ratings, and belief in eyewitness, and a logistic regression for verdict. Descriptives are listed in Table 1.

Compared to students, community members had higher comprehension concerning exposure duration, $F(1, 424) = 7.78, p = .006, \eta^2_p = .02, d = .27, 95\% \text{ CI } [.08, .45]$, time delay between witnessing a crime and making an identification, $F(1, 424) = 4.91, p = .027, \eta^2_p = .01, d = .23, 95\% \text{ CI } [.04, .41]$, unbiased lineup instructions, $F(1, 423) = 9.06, p = .003, \eta^2_p = .02, d = .30, 95\% \text{ CI } [.11, .48]$, and a lineup procedure (versus a showup), $F(1, 423) = 5.14, p = .024,$
\(\eta_p^2 = .01\) \(d = .20\), 95% CI [.01, .39]. Community members also rated eyewitness confidence higher when police gave confirmatory feedback, \(F (1, 424) = 8.03, p = .005, \eta_p^2 = .02, d = .28\), 95% CI [.09, .47].

Despite these more accurate responses to the *general* impact of witnessing and identification conditions on eyewitness accuracy and confidence, community members rated the eyewitness as more credible than students, \(F (1, 425) = 3.73, p = .05, \eta_p^2 = .01, d = .19\), 95% CI [.01, .38], and reported more belief in her identification, \(F (1, 411) = 3.77, p = .05, \eta_p^2 = .01, d = .18\), 95% CI [.00, .36], regardless of the actual quality of those conditions. In addition, participant type did not influence jurors’ perceptions of how helpful the instructions were for evaluating the eyewitness testimony, \(F (1, 402) = .004, p = .95, \eta_p^2 = .00\). Finally, any differences observed in ratings of identification accuracy, witnesses, or belief in the eyewitness did not translate into differences in convictions, directly or in interaction with the quality of the witnessing and identification conditions or safeguard types (all \(p < .05\)). Therefore, participant type was not included in subsequent analyses.

**Do Jurors have Knowledge of Eyewitness Factors and Do Safeguards Increase Such Knowledge?**

Jurors, on their own, appeared to be knowledgeable concerning the six eyewitness factors assessed, with all means above the mid-point (see Table 1 for all means). To determine if any of the safeguards served as an educational tool concerning various eyewitness factors, I ran a series of one-way ANOVAs to examine the main effect of safeguard on each of the comprehension ratings. I observed a significant main effect of safeguard type on lineup instructions, \(F (4, 445) = 4.50, p = .001\). The Tukey test for mean differences indicated that *Henderson* instructions, \(d = \)
.47, 95% CI [.17, .77], expert testimony, $d = .54$, 95% CI [.24, .83], and research-based instructions, $d = .53$, 95% CI [.23, .82] all had significantly higher means than the control condition. That is, participants in these conditions correctly noted that the likelihood of identification accuracy is generally increased with the use of unbiased lineup instructions. None of the other means differed from one another.

None of the safeguards influenced comprehension of exposure duration, weapon presence, time delay, identification procedure type, or feedback on identification accuracy or confidence, all $F$s (4, 445) < 2.18, $p$s > .05. However, planned comparisons using the Tukey test of mean differences indicated that participants had higher comprehension for weapon presence with expert testimony compared to the control condition, $d = .45$, 95% CI [.16, .75]. None of the other means differed from one another (see Table 1).

**Do Safeguards Sensitize Jurors?**

I ran a series of ANOVAs to test for the main and interactive effects of the manipulated variables on witness ratings, belief in the eyewitness, and belief that the instructions help jurors evaluate eyewitness testimony. I ran a logistic regression for verdict. Descriptives are listed in Table 1. Correlations among all proposed mediators and dependent variables are presented in Table 2. Conditional means are presented in Table 3.

Concerning detective credibility, there was a significant main effect of quality of identification conditions, $F (1, 431) = 24.76$, $p < .001$, $\eta_p^2 = .05$, and safeguard type, $F (4, 431) = 2.44$, $p = .05$, $\eta_p^2 = .02$ (see Table 1). Specifically, participants viewed the detective as more credible when identification conditions were good (vs. poor), $d = .46$, 95% CI [.27, .64]. When an expert testified, participants reduced credibility ratings for the detective compared to the
control condition, $d = .50$, 95% CI [.21, .80]. None of the other means significantly differed from one another. In addition, none of the two or three-way interactions were significant, $F$s (4, 431) < 2.00, $ps > .05$, $\eta^2_p < .02$ (see Table 3).

For eyewitness credibility, there was a main effect of quality of identification conditions, $F (1, 432) = 5.85$, $p = .02$, $\eta^2_p = .01$ (see Table 1). Specifically, participants viewed the eyewitness as more credible when identification conditions were good (vs. poor), $d = .23$, 95% CI [.04, .41]. None of the two or three-way interactions were significant, $F$s (4, 432) < 1.74, $ps > .05$, $\eta^2_p < .02$ (see Table 3). There were no significant main or interactive effects for belief in the eyewitness, $F$s (1-4, 431) < 1.83, $ps > .05$, $\eta^2_p < .02$, or belief that instructions help jurors evaluate eyewitness testimony, $F$s (1-4, 422) < 2.47, $ps > .05$, $\eta^2_p < .01$ (see Tables 1 and 3).

Finally, I regressed the dichotomous verdict on quality of witnessing and identification conditions, safeguard type, and all two and three way interactions involving safeguard type in a logistic regression analysis. Quality of witnessing and identification conditions, and safeguard type were entered in the first step and twelve interactions were entered in the second step. Safeguard type was dummy coded, allowing me to directly compare the impact of each safeguard to the absence of any safeguard as well as its interaction with the quality of witnessing and identification conditions. The omnibus test for the first step with all partial effects was significant, $\chi^2(6, N = 448) = 19.52$, $p = .003$, Nagelkerke $R^2 = .06$. There was a significant effect of quality of identification conditions and expert testimony (vs. no safeguard) on verdict. The odds of a guilty verdict were 1.56 times greater when identification conditions were good (vs. poor), Wald $\chi^2(1, N = 448) = 5.22$, $p = .02$, $Exp(B) = 1.56$, 95% CI [1.07, 2.28]. When jurors
heard expert testimony (vs. no safeguard), the odds of a guilty verdict were .39 times lower, Wald $\chi^2(1, N = 448) = 9.12, p = .003, \text{Exp}(B) = .39, 95\% \text{ CI} [.21, .72].$

I also compared each of the safeguards to one another by switching out the reference group. The odds of conviction were lower with expert testimony compared to Henderson instructions, Wald $\chi^2(1, N = 448) = 8.85, p = .003, \text{Exp}(B) = .39, 95\% \text{ CI} [.21, .73]$, research-based instructions, Wald $\chi^2(1, N = 448) = 4.67, p = .03, \text{Exp}(B) = .51, 95\% \text{ CI} [.28, .94]$, and expert + Henderson, Wald $\chi^2(1, N = 448) = 7.30, p = .01, \text{Exp}(B) = .43, 95\% \text{ CI} [.23, .79]$. With the exception of expert testimony, none of the safeguards differed from one another or the control condition. All interactions were non-significant, Wald $\chi^2(1, N = 448) < 2.24, ps > .05, \text{Exp}(B) < 3.16$.

**Mediation**

I did not observe any interactions in the previous analyses. I did, however, find that the quality of identification conditions was predictive of both detective and eyewitness credibility, as well as verdict. In addition, expert testimony (vs. no safeguard) was predictive of detective credibility and verdict. Therefore, I followed the moderation analyses up with two mediational analyses (PROCESS Model 4; Hayes, 2013) using bias-corrected 10,000 sample bootstrap confidence intervals. The first model examined the potential mediating roles of detective and eyewitness credibility on the relationship between quality of identification conditions and verdict, with safeguard (dummy coded) and quality of witnessing conditions included as covariates (see Table 4 for total, direct, and indirect effects). The second model examined the potential mediating role of detective credibility on the relationship between expert testimony (vs. no safeguard) and verdict, including the remaining dummy-coded safeguards, and quality of
witnessing and identification conditions as covariates (see Table 5 for total, direct, and indirect effects).

In the first model, quality of identification conditions had a positive effect on both detective credibility, \( b = .55, 95\% \text{ CI} [.32, .78] \), and eyewitness credibility, \( b = .29, 95\% \text{ CI} [.05, .53] \). Additionally, eyewitness credibility was a positive predictor of verdict, \( b = 1.09, 95\% \text{ CI} [.84, 1.34] \). However, detective credibility did not predict verdict after accounting for eyewitness credibility, \( b = .05, 95\% \text{ CI} [-.17, .28] \). There was an indirect effect of the quality of identification conditions on verdict through eyewitness credibility, \( b = .32, 95\% \text{ CI} [.04, .61] \), such that good identification conditions increased perceptions of the eyewitness’ credibility which, in turn, increased the odds of a conviction (see Figure 1).

In the second model, expert testimony (vs. no safeguard) had a negative effect on detective credibility, \( b = -.58, 95\% \text{ CI} [-.93, -.22] \). Detective credibility had a positive effect on verdict, \( b = .58, 95\% \text{ CI} [.41, .76] \). There was an indirect effect of expert testimony (vs. no safeguard) on verdict through detective credibility, \( b = -.34, 95\% \text{ CI} [-.60, -.13] \), such that expert testimony reduced perceptions of the detective’s credibility which, in turn, reduced the odds of a conviction (see Figure 2).
Chapter 12: Study 1 Discussion

I tested two key assumptions made by the New Jersey Supreme Court in the *Henderson* decision. First, the Court assumed case-specific instructions would sensitize jurors to the quality of eyewitness identifications. Second, the Court assumed the instructions would be at least as effective as expert testimony in accomplishing this goal. However, I did not observe any significant two- or three-way interactions with the quality of witnessing and identification conditions and safeguard type. These null effects indicate that all forms of the safeguards (*Henderson* instructions, research-based instructions, expert testimony, and expert testimony + *Henderson* instructions) were ineffective in this regard.

The odds of conviction depended on both the quality of identification conditions and presence of expert testimony. Mock jurors, on their own, were rather sophisticated at evaluating police practices by increasing convictions when identification conditions were good and reducing convictions when conditions were poor. Eyewitness credibility mediated this relationship. Specifically, the credibility of the eyewitness hinged on how police performed; when police gave unbiased lineup instructions, conducted a lineup procedure (instead of a showup), and withheld confirmatory feedback from the eyewitness after she identified the defendant, jurors increased their ratings of the eyewitness’ credibility and, in turn, increased convictions. A few studies have observed juror sensitivity to system variables, such as lineup instructions (Cutler et al., 1988) and foil-biased lineups (Devenport et al., 2002) without any safeguards. I believe this is the first study to demonstrate juror sensitivity to identification procedure (lineup vs. showup) and confirmatory feedback, though the confounding of these factors indicates additional studies need to be conducted to tease apart findings.
The addition of expert testimony did not increase juror sensitivity. Relative to all of the other safeguard conditions (Henderson, research-based instructions, expert + Henderson instructions, and no safeguard), expert testimony induced skepticism about the accuracy of the eyewitness identification, such that convictions were reduced regardless of the quality of witnessing and identification conditions. This skepticism effect replicates previous findings (Cutler et al., 1990). Detective credibility mediated the relationship between expert testimony and verdict such that those exposed to expert testimony reduced credibility ratings for the detective which, in turn, decreased convictions. The expert caused jurors to become more critical of how the police handled this case.

Henderson instructions, research-based instructions, expert testimony + Henderson instructions, and the quality of witnessing conditions did not influence verdict decisions. Thus, my second hypothesis that research-based Henderson instructions and expert testimony would be equally effective was not supported. The lack of any effect of eyewitness instructions on verdict supports previous research examining the Telfaire instructions (Cutler et al., 1990; see also Greene, 1988 Experiment 1), but stands in contrast to the skepticism effects found with the Henderson instructions in Berman et al. (2015). The null findings concerning the quality of witnessing conditions suggest such conditions are particularly difficult for jurors to evaluate and adjust their verdict decisions accordingly (see Desmarais & Read, 2011). As noted previously, none of the proposed safeguards assisted jurors in evaluating this type of evidence. To my knowledge, no eyewitness instruction studies have observed sensitivity effects for estimator variables (cf., Ramirez et al, 1996, though even they cannot say with certainty that jurors were
sensitized to estimator variables with eyewitness instructions because they confounded system and estimator variables in one manipulation).

Despite little influence on verdict, the safeguards improved jurors’ understanding of some system and estimator variables. Jurors who received *Henderson* instructions, expert testimony, or research-based instructions (compared to no safeguard) better understood the impact of lineup instructions on identification accuracy. This is noteworthy given that the biggest difference in understanding between jurors and experts concerns lineup instructions (41% vs. 98% agreement that “lineup instructions can affect an eyewitness’ willingness to make an identification”; Benton et al., 2006). Furthermore, those who heard expert testimony had higher comprehension of weapon presence (compared to no safeguard).

Across safeguard conditions, however, eyewitness knowledge was above the scales’ midpoint for all six system and estimator variables, suggesting jurors are not oblivious to factors influencing identification accuracy. The current findings indicate that jurors, without the instructions, were able to incorporate system variable knowledge, but not estimator variable knowledge, into their decision making. Perhaps this is why jurors did not rate the *Henderson* instructions as more helpful for evaluating the eyewitness testimony than the control or expert testimony conditions (where jurors would have only heard instructions on reasonable doubt and burden of proof). Thus, at least for identification conditions, the instructions were not necessary. However, the same cannot be said for witnessing conditions. Thus, jurors may first need to endorse the usefulness of a safeguard for evaluating witnessing conditions before they can appropriately adjust their verdict decisions (see Leippe & Eisenstadt, 2009).
Broken down by participant type, community members were more knowledgeable than students on five out of six system and estimator variables. The only factor community members did not understand better than students was weapon focus. However, community members (compared to students) believed the eyewitness was more credible and accurate regardless of the actual witnessing and identification conditions. Additionally, I did not observe any differences in how helpful the instructions were for evaluating eyewitness testimony or convictions as a function of participant type. Thus, any superior knowledge that community members may have had over students was not integrated into their decision making. The courts’ hesitation to rely on studies conducted solely with college students may be overly cautious as verdict decisions, at least in this study, did not differ between these samples (see also Bornstein, 1999).

Limitations

The current study suffered from a few limitations. First, proposed mediators were assessed after participants rendered a verdict. It is possible that determining a verdict first shaped subsequent responses to the proposed mediators. Second, I worked on the assumption that judges will only include instructions on factors considered problematic in a case. Therefore, the instructions given under varying witnessing and identification conditions were not equivalent. I suspect determining which factors are problematic will be a difficult task for judges to determine in actual cases, especially for factors that can only be estimated after the crime (i.e., estimator variables). The New Jersey Supreme acknowledged this difficulty themselves when they decided that an identification’s admissibility should not hinge on estimator variables alone.

…We anticipate that eyewitness identification evidence will likely not be ruled inadmissible at pretrial hearings solely on account of estimator variables. For example, it
is difficult to imagine that a trial judge would preclude a witness from testifying because the lighting was “too dark,” the witness was “too distracted” by the presence of a weapon, or he or she was under “too much” stress while making an observation. How dark is too dark as a matter of law? How much is too much? What guideposts would a trial judge use in making those judgment calls? (*New Jersey v. Henderson*, 2011, p. 994).

These same questions concern the administration of instructions given to jurors. For simplicity purposes, I decided to include instructions only on factors where research has clearly established that reliability is reduced (e.g., 10 second exposure duration vs. 45 seconds), which also follows the format typically found with experts who testify for the defense. Of course, it is possible that delivering instructions in good witnessing conditions could produce a sensitizing effect reflected in increased convictions.

**Conclusions**

As a leader in criminal law, the New Jersey Supreme Court’s decision to implement case-specific eyewitness instructions may encourage other states to follow suit. The widespread adoption of instructions similar to *Henderson* should depend on its effectiveness. Unfortunately, New Jersey and other states considering adopting *Henderson* style eyewitness instructions should give pause as the current results suggest these instructions have little impact on jurors’ decisions. Overall, it appears the *Henderson* Court overestimated the ability of these case-specific eyewitness instructions to assist jurors to better evaluate eyewitness evidence or to lessen the need for expert testimony.

The finding that jurors adjusted their verdicts based on the quality of some system variables suggests they may not need any safeguards to help them better evaluate the quality of
identification conditions as it concerns lineup instructions, identification procedure, and/or confirmatory feedback. Though just how large an effect on verdicts the courts may desire is an open question that will merit separate investigation. However, I did not detect this same effect on verdict for estimator variables, suggesting jurors are not sensitive on their own to the quality of witnessing conditions. Therefore, safeguards could provide some assistance to jurors to evaluate this type of eyewitness evidence. Unfortunately, none of the safeguards tested in this study was able to induce sensitivity to estimator variables.

**Research Implications**

Since none of the variations of Henderson instructions sensitized jurors (and courts prefer instructions over expert testimony), I examined additional ways to modify and improve the effectiveness of Henderson instructions relying, in part, on the I-I-Eye (Pawlenko et al., 2013). Replication of the Pawlenko et al. (2013) study is necessary as results may have been due to how the information was organized and administered to participants. For example, the I-I-Eye is repetitive, though substantially shorter than corresponding Henderson instructions (1389 vs. 2345 words). However, the I-I-Eye reviews a number of factors which may or may not be relevant to a particular case, while the New Jersey Supreme Court intended Henderson instructions to be case-specific. Thus, Henderson may be much shorter depending on the factors the judge decides to include, but may not address all potential factors. The sensitivity effects with the I-I-Eye teaching aid (Pawlenko et al., 2013) may, therefore, be due to its general, blunderbuss (vs. case-specific) nature that educates jurors on a variety of factors that can serve to either increase or decrease the reliability of a specific identification. Giving the same set of instructions to jurors regardless of the witnessing and identification conditions associated with a
specific identification eliminates the problems outlined in the previous study’s limitations and removes the discretion allotted to judges.

Juror sensitivity with the I-I-Eye may also be the result of prompts within the instruction, pressing jurors to evaluate the evidence before determining a verdict (e.g., for how long did the eyewitness view the perpetrator?). The results of the first study indicate that thorough evidence evaluation did not appear to occur with mock jurors who heard the Henderson instructions (see also Berman et al., 2015). As such, simply instructing jurors to consider a particular factor is insufficient. The addition of prompts may improve the effectiveness of Henderson instructions, though it is unknown whether such modifications could increase sensitivity to estimator variables because these were not manipulated in Pawlenko et al. (2013).

One possible reason for the ineffectiveness of the original Henderson instructions may be that jurors believed the instructions were too biased for the defense, leading them to be dismissive of the instructions entirely. Prompts in addition to providing information on a variety of non-case specific factors may neutralize the instructions, allowing jurors to better notice when certain factors either weaken or strengthen identification reliability.

Another possible avenue for improving juror decisions and ability to evaluate complex evidence is through the use of special verdicts or interrogatories. More common in civil than criminal trials, jurors may be required to answer a series of factual questions with a general verdict (called a general verdict with interrogatories) or without a general verdict (called a special verdict; see Wiggins & Breckler, 1990). These additional questions may provide jurors with a framework to review the overall quality of the evidence and ease determination of a final verdict (Wiggins & Breckler, 1990). Jurors who served in actual criminal and civil trials reported
that special verdict forms were helpful for evaluating large amounts of information (Heuer & Penrod, 1994). In addition, these jurors self-reported higher comprehension of the judge’s instructions, and more satisfaction and confidence in their verdict with special verdict forms, though the type of verdict form was not manipulated. When mock jurors were randomly assigned to render either a general or special verdict, comprehension of burden of proof was higher among those who rendered special verdicts, though there were no differences in trial outcome (Wiggins & Breckler, 1990). Overall, little research has been conducted on alternative verdict forms, especially in conjunction with eyewitness instructions, leaving its impact on juror decisions unknown.

Broader research on cognitive dissonance theory suggests that once jurors determine a verdict, they will align subsequent responses with the initial verdict decision to avoid any discomfort from conflicting decisions (i.e., dissonance reduction; Festinger, 1957). That is, jurors may distort their impressions of witnesses or the quality of an eyewitness identification to support their verdict preference. However, less is known about the impact of having jurors evaluate evidence before determining the verdict. Therefore, I conducted a follow up study that examined whether incorporating features of the I-I-Eye into the Henderson instructions along with pre-verdict evidence evaluation (i.e., interrogatories) could induce juror sensitivity. I hypothesized that instructions would be most effective when jurors are forced to evaluate the evidence prior to determining a verdict compared to afterward.
Participants

A total of 482 jury-eligible community members were recruited via mTurk.com. To ensure high quality data, only those with 95% approval ratings were allowed to participate (Peer, Vosgerau, & Acquisti, 2014). Participants ranged in age from 18-79 ($M=37.11$, $SD=12.31$). Forty-four percent of participants were female and the majority were White (80.5%), followed by African American (8.3%), Asian (4.8%), Hispanic (3.3%), and Other (3.1%). Finally, 51.2% of participants indicated they were employed full-time, 16.2% part-time, 14.3% unemployed, 6.6% students, and the remaining listed retired or other.

Design and Materials

A 3 (Instruction type: none vs. original Henderson vs. modified Henderson) x 2 (Evidence strength: weak vs. strong) x 2 (Interrogatories timing: pre vs. post-verdict) factorial design was implemented.

Trial stimulus. Mock jurors read a criminal trial transcript involving an eyewitness identification. The defendant was charged with robbery. All participants read New York standard judicial instructions, opening and closing statements from both attorneys, and direct and cross examination of the detective, eyewitness, and an alibi witness. The presence of Henderson instructions (original or modified) varied as well as evidence strength concerning the eyewitness identification.

Independent Variables

Instruction type. Participants in the control condition only received burden of proof and reasonable doubt instructions. In addition to the burden of proof and reasonable doubt
instructions, participants in the original and modified Henderson instructions conditions read about a total of 15 factors known to influence eyewitness accuracy (i.e., stress, duration, weapon presence, time elapsed, distance, lighting, intoxication, disguise, prior perpetrator description, confidence and accuracy, lineup composition, fillers, double blind administration, lineup instructions, and confirmatory feedback; see Appendix C). These factors were not specific to the case they read. The modified version of the instructions condensed the beginning and end of the original instructions, which outlined the stages of memory and reasonable doubt (which was already covered separately in all conditions), but left the explanation of each specific factor intact. For example, jurors in both the original and modified versions of Henderson read about duration:

Duration – The amount of time the eyewitness has to observe an event may affect the reliability of an identification. Although there is no minimum time required to make an accurate identification, a brief or fleeting contact is less likely to produce an accurate identification than are more prolonged exposure to the perpetrator. In addition, time estimates given by witnesses may not always be accurate because witnesses tend to think events last longer than they actually did.

But in addition, those assigned to modified Henderson instructions received this prompt at the end of each factor:

Now Ask Yourself: For how long did the eyewitness view the perpetrator? Does this enhance or impair his/her accuracy?

To contrast with the I-I-Eye, here is an excerpt on duration (Pawlenko et al., 2013):

For how long did the eyewitness view the perpetrator? The less time an eyewitness has to witness a crime the less information they will remember about it.

Evidence strength. The quality of the witnessing conditions was manipulated. In the weak condition, participants read about a total of three estimator variables that were problematic in the current case. Specifically, the perpetrator used a knife, the eyewitness only viewed the
perpetrator’s face for 10 seconds, and there was a one month delay between when the crime occurred and the identification was made. In strong condition, the perpetrator did not have a weapon, the eyewitness viewed the perpetrator’s face for 45 seconds, and there was only a one day delay between when the crime occurred and the identification was made. In all conditions, the officer conducted a fair double-blind six person lineup, gave unbiased lineup instructions, and refrained from giving confirmatory feedback to the eyewitness.

**Interrogatories timing.** Participants were asked a variety of questions about the eyewitness evidence either before or after they rendered a verdict (see belief in the eyewitness and comprehension ratings below).

**Manipulation Checks**

Participants completed a total of five manipulation checks. Mock jurors indicated the time delay between witnessing the crime and making an identification, the amount of time the victim had to see the perpetrator’s face, and whether the perpetrator used a weapon. Mock jurors also indicated whether the judge gave instructions on eyewitness identification issues.

**Mediators**

**Belief in eyewitness.** Participants provided the probability that the eyewitness correctly identified the defendant as the perpetrator (0% = not at all probable, 100% = completely probable).

**Comprehension ratings.** Mock jurors indicated whether each of the three manipulated estimator variables (i.e., weapon, duration, and delay) influenced the likelihood of an accurate identification *in this case* using nine-point Likert scales (1 = Strongly Reduces Accuracy, 9 = Strongly Increases Accuracy). Items were recoded as necessary so that higher values indicated
higher comprehension and averaged to create a single scale measure ($\alpha = .86$). Participants were also asked to indicate how much of an influence the behavior of the police had on the accuracy of the eyewitness’ identification (1 = Strongly Reduces Accuracy, 9 = Strongly Increases Accuracy).

**Witness ratings.** Participants rated their impressions of the detective and eyewitness on a series of seven-point bipolar adjective pairs. These items included trustworthy-untrustworthy, convincing-unconvincing, and certain-uncertain. For each scale, I recoded items as necessary so that higher values represented more favorable ratings and averaged the items to create a single scale measure (detective, $\alpha = .90$; eyewitness, $\alpha = .91$).

**Dependent Measures**

**Verdict.** Mock jurors provided dichotomous judgments of guilt (0=Not guilty; 1=Guilty).

**Instructions help.** Mock jurors rated the extent to which they agreed the judge’s instructions helped them to evaluate the eyewitness testimony using nine-point Likert scales (1 = Strongly Disagree, 9 = Strongly Agree).

**Procedure**

Community members were recruited via mTurk.com in exchange for $1. Participants were pre--screened to ensure they were jury eligible (i.e., 18 years of age or older and a US citizen). After participants provided informed consent, they completed the voir dire questionnaire. Participants were randomly assigned to one of 12 conditions. Participants then read the trial transcript. The presence of Henderson instructions varied as well as the quality of the eyewitness identification. Next, participants completed the post-trial questionnaire where the timing of comprehension ratings also varied. Specifically, participants determined a verdict and
witness ratings either before or after completing comprehension ratings and belief in the eyewitness. Manipulation checks were always completed last. After completing the questionnaire, participants were debriefed. The entire study took approximately 45 minutes.

**Data Analytic Plan**

I conducted $\chi^2$ analyses to examine manipulation checks. To establish if there were main or interactive effects of instruction type, evidence strength, and interrogatories timing on the various mediators, I ran a series of ANOVAs and a logistic regression for verdict (followed by PROCESS Model 2 to interpret interactions; Hayes, 2013). This was followed up with a moderated mediation analysis (PROCESS Model 10; Hayes, 2013) using bias-corrected 10,000 sample bootstrap confidence intervals. This model allowed me to examine whether any conditional effects directly or indirectly influenced verdict. I report unstandardized regression coefficients for the PROCESS models as recommended by Hayes (2013).
Chapter 14: Study 2 Results

Manipulation Checks

Participants were successful at noticing the evidence strength manipulation. Compared to those in the weak condition, participants in the strong condition were more likely to respond that the time delay between witnessing the crime and making the identification was only one day (7.9% vs. 93.6%), $\chi^2(1, N = 480) = 352.25, p < .001, \varphi = -.86$; that the eyewitness viewed the perpetrator for about 45 seconds (3.7% vs. 95.1%), $\chi^2(1, N = 481) = 400.15, p < .001, \varphi = .91$; and that the perpetrator did not have a weapon (3.3% vs. 92.1%), $\chi^2(1, N = 480) = 375.29, p < .001, \varphi = -.88$.

Mock jurors were more likely to respond that the judge discussed identification procedures when they heard the original (80.6%) or modified (65.8%) Henderson instructions than the control condition (13.6%), $\chi^2(2, N = 482) = 158.44, p < .001, \varphi = .57$. They were also more likely to indicate that the judge discussed witnessing conditions when they heard the original (90.3%) or modified (83.2%) Henderson instructions compared to the control condition (18.6%), $\chi^2(2, N = 482) = 215.95, p < .001, \varphi = .67$. These findings support those of Peer et al. (2014) who found high reputation mTurkers (i.e., 95% approval ratings) produce high quality data. Additionally, analyses were conducted with the full sample and compared to analyses without participants who failed more than three of the five manipulation checks (n=22). Results were the same across the different analyses. Below, I report analyses with all participants.

Sensitivity Effects: Moderation analyses

Descriptives are listed in Table 6. Correlations among all proposed mediators and verdict are presented in Table 7. For participants’ ratings of how helpful the instructions were for
evaluating eyewitness testimony, there was a significant main effect for instruction type, such that jurors believed both the original and modified Henderson instructions were more helpful for evaluating eyewitness testimony compared to the control condition, $F(1, 470) = 11.08, p = .001, \eta^2_p = .05$ (original: $d = .44$, 95% CI [.19, .69]; modified: $d = .43$, 95% CI [.22, .63]). This effect was qualified by a significant interaction between evidence strength and instruction type, $F(2, 470) = 4.35, p = .01, \eta^2_p = .01$ (see Table 8 for means). Using a Bonferroni correction, there was a conditional effect of evidence strength for the original Henderson instruction condition, $F(1, 470) = 5.99, p = .01, \eta^2_p = .01$. Jurors rated the instructions as more helpful for evaluating the eyewitness testimony when evidence was strong compared to weak, $d = .58$, 95% CI [.17, 1.00].

For participants’ belief in the eyewitness, there was a significant interaction between evidence strength and instruction type, $F(2, 462) = 5.04, p = .01, \eta^2_p = .02$ (see Table 8 for means). Using a Bonferroni correction, there was a conditional effect of evidence strength for the control condition, $F(1, 462) = 20.28, p < .001, \eta^2_p = .04$, and the modified Henderson instructions condition, $F(1, 462) = 28.71, p < .001, \eta^2_p = .06$. Jurors were able to distinguish between the strength of evidence and adjust their perceptions of the quality of the eyewitness identification accordingly when they did not hear any eyewitness instructions, $d = .64$, 95% CI [.35, .93], and even more so when they heard modified Henderson instructions, $d = .78$, 95% CI [.48, 1.08]. The three way interaction was non-significant, $F(2, 470) = .15, p > .05, \eta^2_p = .001$, as was the main effect of interrogatories timing, $F(1, 470) = .51, p > .05, \eta^2_p = .001$.

For comprehension ratings, there was a main effect of evidence strength, $F(1, 469) = 254.37, p < .001, \eta^2_p = .35$. Jurors reported higher comprehension ratings regarding the quality of the witnessing conditions when evidence was strong (versus weak; $d = 1.58$, 95% CI [1.44,
Main effects for interrogatories timing, $F(1, 469) = 1.39, p > .05, \eta^2_p = .00$, and instruction type, $F(2, 469) = 2.21, p > .05, \eta^2_p = .01$, were non-significant as were all two and three way interactions, $Fs(2, 470) < 1.28, p > .05, \eta^2_p s < .01$. In addition, no significant effects were found for the influence of police behavior on the likelihood of an accurate identification, $Fs(2, 462) < 2.45, p > .05, \eta^2_p s < .01$. This is not surprising given the quality of system variables were held constant across conditions.

For detective ratings, there was a main effect of interrogatories timing, $F(1, 470) = 4.01, p = .05, \eta^2_p = .00$, such that jurors rated the detective more favorably when they determined the verdict before evaluating the evidence. However, this effect was weak and the confidence interval crossed zero, $d = .13, 95\% \text{ CI } [-.05, .31]$. Main effects for evidence strength, $F(1, 470) = .04, p > .05, \eta^2_p = .00$, and instruction type, $F(2, 470) = 1.93, p > .05, \eta^2_p = .01$, were non-significant as well as all two and three way interactions, $Fs(2, 470) < 2.30, p > .05, \eta^2_p s < .01$.

For eyewitness ratings, there was a main effect of evidence strength, $F(1, 470) = 6.57, p = .01, \eta^2_p = .01$. Jurors rated the eyewitness more favorably when the quality of the witnessing conditions was good, $d = .31, 95\% \text{ CI } [.18, .43]$. Main effects for interrogatories timing, $F(1, 470) = .05, p > .05, \eta^2_p = .00$, and instruction type, $F(2, 470) = .15, p > .05, \eta^2_p = .00$, were non-significant. In addition, all two and three way interactions were not significant, $Fs(2, 470) < 1.44, p > .05, \eta^2_p s < .01$, though planned comparisons indicated there were conditional effects of evidence strength for the control and modified Henderson conditions (with means in the same direction as the interaction found for belief in the eyewitness; see Table 8 for means).

For verdict, I entered evidence strength, interrogatories timing, and dummy coded instruction type in the first step and all two and three way interactions involving safeguard type
in the second step of a logistic regression. I dummy coded safeguard type and made the original Henderson instructions condition the reference group.\(^2\) The omnibus test for the first step with all partial effects was significant, \(\chi^2(4, N = 482) = 30.19, p < .001\), Nagelkerke \(R^2 = .08\), as was the second step with all interactions, \(\chi^2(8, N = 482) = 14.12, p = .01\), Nagelkerke \(R^2 = .12\). No three way interactions were significant, \(ps > .05\). However, there were significant interactions between evidence strength and the control condition (vs. original Henderson instructions), Wald \(\chi^2(1, N = 482) = 5.41, p = .02\), \(Exp(B) = 3.82\), 95% CI [1.24, 11.84], and between evidence strength and the modified Henderson instructions (vs. original Henderson instructions), Wald \(\chi^2(1, N = 482) = 4.64, p = .03\), \(Exp(B) = 3.37\), 95% CI [1.12, 10.17].

To interpret the significant interactions, I followed this analysis up with PROCESS Model 2 (Hayes, 2013). There was a conditional and positive effect of evidence strength on verdict among those who did not receive any instructions, \(b = .88\), 95% CI [.30, 1.46], and a stronger effect among those who received the modified Henderson instructions, \(b = 1.29\), 95% CI [.69, 1.90]. There was no such relationship for those who received the original Henderson instructions (and was actually in the opposite direction), \(b = -.29\), 95% CI [-1.13, .55]. Jurors were able to distinguish between the quality of evidence on their own, but even more so when they received modified Henderson instructions.

**Moderated Mediation**

None of the three way interactions were significant, nor did I observe any two or three-way interactions between the manipulated variables for comprehension or witness ratings. Therefore, I ran a moderated mediational model (PROCESS Model 10; Hayes, 2013) to determine if the effect of the two-way interaction between evidence strength and instruction type
on verdict operated through belief in the eyewitness. As with the logistic regression, I dummy coded instruction type and made the original *Henderson* instructions condition the reference group.

Again, I found that the interaction between evidence strength and modified *Henderson* instructions (vs. original *Henderson* instructions) was significant, $b = 19.64$, 95% CI [6.85, 32.43], as was the interaction between evidence strength and the control condition (vs. original *Henderson* instructions; $b = 16.19$, 95% CI [3.49, 28.89]) for belief in the eyewitness. In turn, belief in the eyewitness was a significant predictor of verdict, $b = .11$, 95% CI [.09, .13]. The indirect effect of evidence strength on verdict through belief in the eyewitness was conditioned on instruction type (see Figure 3). This indirect effect was significant and positive with modified *Henderson* instructions, $b = 2.11$, 95% CI [1.27, 3.08], but also without any instructions, $b = 1.74$, 95% CI [.93, 2.60]. Since the confidence intervals overlap, the indirect effects of evidence strength on verdict through belief in the eyewitness do not appear to differ for those in the control condition and those who received modified *Henderson* instructions. The conditional indirect effect was not significantly different from zero for those who received the original *Henderson* instructions, $b = .01$, 95% CI [-1.18, 1.12]. Finally, none of the conditional direct effects were significant (control: $b = -.36$, 95% CI [-1.29, .56]; modified: $b = .46$, 95% CI [-.47, 1.40]; original: $b = -.64$, 95% CI [-1.87, .59]), indicating that belief in the eyewitness accounted for the moderating effect of safeguard type on the relationship between witnessing conditions quality and verdict.
Chapter 15: Study 2 Discussion

The findings of the current study indicate jurors on their own were able to evaluate evidence strength. Introducing the original *Henderson* instructions actually served to desensitize jurors, eliminating differences in convictions according to the quality of witnessing conditions. However, modifying the content of the original *Henderson* instructions improved their effectiveness. I observed the biggest differences in both belief in the eyewitness and convictions between weak and strong evidence among those exposed to the modified *Henderson* instructions. This finding complements those of the I-I-Eye which induced sensitivity to identification conditions (see Pawlenko et al., 2013).

While not what I had hypothesized, there were a number of main effects on the outcome measures. Evidence strength influenced jurors' eyewitness and comprehension ratings. Specifically, participants had higher eyewitness credibility ratings and comprehension of estimator variables when evidence was strong. This suggests some sensitivity on the part of jurors who recognized that the credibility of the eyewitness hinged on the quality of the witnessing conditions. Yet it also suggests that jurors may have a particularly difficult time understanding eyewitness factors when they *undermine* reliability. Additionally, participants believed the detective was more credible when they determined the verdict before evaluating the evidence. It is possible that jurors became more critical of the detective’s behavior (regardless of what he actually did) when they evaluated the evidence before determining a verdict. Finally, jurors reported how helpful the instructions were for evaluating the eyewitness testimony. In the control and expert conditions, jurors only received instructions on reasonable doubt and burden of proof, whereas jurors in the remaining conditions received these instructions plus some
variation of *Henderson* eyewitness instructions. I suspected jurors would believe the eyewitness instructions were more helpful than the reasonable doubt instructions alone, yet I observed a peculiar finding. There were only differences in this belief when jurors heard the original *Henderson* instructions, with jurors reporting the original instructions were more helpful when the evidence was strong. Despite believing the instructions were more helpful, comprehension of eyewitness factors did not differ from the control condition. Furthermore, the original *Henderson* instructions served to eliminate the sensitivity jurors displayed on their own and suggests jurors were not aware of the confusion the instructions actually induced!

Unlike Wiggins and Breckler (1990), I did not observe improvements in comprehension when jurors were required to evaluate the evidence prior to determining a verdict. This manipulation also failed to influence verdict decisions, which is in line with Wiggins and Breckler’s findings. Timing of interrogatories was not influential in combination with any other manipulations, including the hypothesized three way interaction between evidence strength, instruction type, and interrogatories timing for any of the proposed mediators or verdict.

Despite no three way interactions, I did observe a significant two way interaction between evidence strength and instruction type. The lowest rate of convictions occurred when evidence was weak with the modified *Henderson* instructions relative to both the control and original *Henderson* instruction conditions. This finding suggests the modified version of *Henderson* is causing jurors to question the reliability of eyewitness identifications precisely when it is most needed. When evidence was strong, convictions increased and were similar among the control and modified *Henderson* instruction conditions. Thus, jurors on their own were better at evaluating the eyewitness evidence than jurors who received the original
Henderson instructions, but it was the modified instructions that sensitized jurors the most. The effect on verdict can be explained by changes in belief in the eyewitness. Participants in the modified instructions and control conditions increased their belief in the eyewitness, and in turn, increased convictions when evidence was strong relative to weak. I did not observe a similar conditional indirect effect for those in the original Henderson instructions condition, though this is not surprising given that convictions and belief in the eyewitness among participants in this condition did not differ according to the strength of evidence.

Overall, the original Henderson instructions had no impact on comprehension or witness ratings, belief in the eyewitness, or verdict, alone or in combination with evidence strength or interrogatories timing. This null finding also occurred in Study 1 and suggests changes to the instructions must be made to assist jurors in evaluating eyewitness evidence. With this in mind, I shortened the original instructions, did not tailor them to the facts of the case (i.e., made the instructions general in nature), and added prompts directing jurors to determine whether each factor was a problem in the current case. These changes improved the ability of jurors to evaluate the eyewitness evidence.

I should note that both the original and modified instructions were not case-specific. While the Henderson Court endeavored to tailor case specific instructions, judges may have difficulty figuring out which factors are issues in a specific case and which are not. For example, how much stress should the eyewitness express suffering before judges deem the stress instruction necessary? What is a sufficient delay between witnessing a crime and making an identification that warrants having the time elapsed instruction included? It is also unclear whether instructions on factors that increase the reliability of the identification might be included
as well. Thus, relying on a general instruction, such as the modified one used in the second study, eliminates the subjective nature of the original *Henderson* instructions and increases the fairness of the procedure for both the State and the defense.

**Limitations**

The second study is not without limitations. First, I utilized a convenience sample of mTurkers. However, few differences exist in work quality between online and in-person participants (Casler, Bickel, & Hackett, 2013). In addition, I restricted the sample to those with 95% approval ratings to increase data quality (Peer et al., 2014). I also used an abbreviated trial transcript of a robbery. Yet, ecological validity does not appear to influence outcomes in legal contexts (Bornstein, 1999).

**Conclusions**

While the New Jersey Supreme Court intended the original *Henderson* instructions to assist jurors in evaluating eyewitness identification evidence, I observed little impact on a variety of measures, including belief in the eyewitness and verdict. Others have found that these instructions cause jurors to become overly critical of the evidence (see Berman et al., 2015; Yoakum & Papailiou, 2014). In contrast, the I-I-Eye causes jurors to become sensitive to the quality of system variables and adjust their verdicts accordingly (Pawlenko et al., 2013). Yet, its format and delivery (i.e., PowerPoint slides) is quite different from that of traditional judicial instructions. I sought, therefore, to create a modified version of *Henderson* that incorporated key features of the I-I-Eye while staying true to the format of the original *Henderson* instructions. These instructions could be delivered in the same way as typical instructions (verbally by the judge and/or via written instructions), making implementation easy.
The results of the second study suggest the use of interrogatories (or a special verdict) do not influence jurors’ comprehension or witness ratings, or their verdict decision. Additional research is needed on its use, but it appears jurors, at least in this study, are relying on the evidence in interaction with modified instructions (and to a lesser extent, on their own) to determine both their belief in the eyewitness and verdict. Because several features of the I-I-Eye (Pawlenko et al., 2013) were implemented in this second study, I am unable to pinpoint what caused the change in juror sensitivity. It is possible that the concise nature of the instructions may have helped jurors focus on the key factors that were problematic in the case and adjust their verdicts according to the impact of those factors. Or perhaps the prompts after each factor induced jurors to reflect on the quality of the eyewitness identification and create a crude checklist of factors that increased or hindered accuracy. It is also possible that the combination of changes is what led to sensitivity.

While it is premature to recommend courts use the simplified, general eyewitness instructions created for the second study, these results indicate the modified instructions are an improvement over the original *Henderson* instructions and provide a somewhat promising avenue for reducing wrongful convictions due to mistaken identifications. Replication is necessary to establish that jurors’ verdicts are reflective of the quality of evidence with the use of modified instructions. Furthermore, it is unclear whether these same modified instructions can induce sensitivity to system and other estimator variables. As Chief Justice Rabner indicated, “Only with a fully informed and properly instructed jury can justice be served” (New Jersey Supreme Court, 2012b).
Chapter 16: Conclusions and Future Directions

The results of the current studies add to our understanding concerning the effectiveness of the newly implemented eyewitness judicial instructions in New Jersey. The New Jersey Supreme Court made two key assumptions that were empirically tested. First, the Court assumed these case-specific instructions would sensitize jurors to the quality of eyewitness identifications. Second, the Court assumed the instructions would be at least as effective as expert testimony in accomplishing this goal.

In the first study, I tested the absolute and relative effectiveness of various forms of judicial instructions and expert testimony. The instruction variations included the original Henderson instructions, a research enhanced version that included meta-analyses to further inform jurors of the magnitude of each eyewitness factor on identification reliability, and a combination of both the original Henderson instructions and expert testimony. All forms of instructions and expert testimony were tailored to the facts of the case as envisioned by the New Jersey Supreme Court.

In contrast to the Court’s hypotheses, I observed no effects for any variation of Henderson instructions nor any two- or three-way interactions with the quality of witnessing and identification conditions. Thus all forms of Henderson in Study 1 failed to assist jurors in evaluating a variety of system and estimator variables in a specific identification. This may be due to lack of improved comprehension for eyewitness identification factors (cf., lineup instructions) or a lack of belief that the instructions helped jurors evaluate the eyewitness testimony, which may be prerequisites for changes in beliefs for specific eyewitnesses and/or verdict decisions (see Leippe & Eisenstadt, 2009).
While instructions had no impact on juror decisions, expert testimony did, albeit not in the anticipated manner. That is, I hypothesized expert testimony would induce sensitivity to system and estimator variables, supporting past research (Cutler et al., 1989). Instead, the presence of expert testimony induced skepticism regardless of the quality of witnessing and identification conditions. Expert testimony reduced perceptions of detective credibility, which in turn reduced convictions relative to all of the other safeguards as well as no safeguard at all. This is not entirely surprising given past research (see Martire & Kemp, 2011) and the fact that when witnessing and identification conditions were good, the safeguards included little information to guide jurors' decisions. The information that was included was still critical of the identification. This, of course, was by design given the focus has been on drawing attention to factors that decrease reliability, but not necessary factors that increase reliability. Thus, it is possible that the Henderson Court desired skepticism when evidence was good, and even more skepticism when the evidence was poor. Whatever their preference, jurors assigned to expert testimony did not distinguish between good and poor witnessing and identification conditions. Jurors may have relied on the simple presence of the expert testimony in these conditions as a heuristic cue instead of attending to the evidence itself (Chaiken, 1980).

Another relevant finding to the first study indicated that jurors were sensitive, on their own, to the quality of identification conditions, but not witnessing conditions. As a result of this finding and the ineffectiveness of multiple forms of Henderson, I decided to conduct a follow up study to examine whether additional modifications to Henderson could induce sensitivity for estimator variables, as research had yet to establish this possibility.
In the second study, I did not make the instructions case-specific. Instead, jurors in all conditions heard instructions on a total of 15 factors that may or may not have been problematic for the identification in question. Furthermore, the modified instructions were condensed down to the core factors, eliminating explanations on the stages of memory and reasonable doubt (which was already included in a separate instruction in all conditions). I also included prompts after the explanation of each factor that rhetorically asked jurors to determine the impact of each factor on identification reliability. These modifications were reflective of the I-I-Eye teaching aid (Pawlenko et al., 2013), which induced sensitivity to system variables previously, though these researchers did not manipulate estimator variables.

Similar to Study 1, the original *Henderson* instructions had no impact on outcome measures. However, the modified version of the *Henderson* instructions induced sensitivity to witnessing conditions. The lowest rate of convictions occurred in the weak evidence strength condition for those who received the modified instructions. Thus, these initial findings suggest that the *Henderson* Court may accomplish their goal of drawing jurors’ attention to problematic identifications with an instruction that varies only slightly from the one they originally commissioned.

Interestingly, jurors reported that the original *Henderson* instructions helped them to evaluate eyewitness testimony, particularly when the evidence was strong (compared to weak). Yet, neither the original or modified instructions increased comprehension compared to the control condition. Thus, more work must be done to establish safeguards that can accomplish the goals of both improving knowledge and increasing sensitivity reflected in verdict decisions (see Leippe & Eisenstadt, 2009).
The current studies are among the first to test the newly formed case-specific instructions proposed by New Jersey. Previous versions of eyewitness instructions have included a fixed number of eyewitness factors, which were not based on psychological research, and failed to provide an adequate explanation for how each factor influences identification accuracy (e.g., Telfaire, see Greene, 1988). While the Henderson instructions addressed these shortcomings, these studies suggest they are not sufficient to induce sensitivity to evidence strength. However, modifying the original format of the instruction by condensing its length, drawing attention to a variety of factors that could decrease or increase reliability, and asking jurors to reflect on the impact of each factor for the current identification appears to improve its sensitizing effect.

**Future Directions**

There are many unanswered questions concerning eyewitness identifications and courtroom safeguards. First, in neither study were jurors sensitive to witnessing conditions on their own. One possible explanation may be due to the nature of the crimes. The literature on eyewitness identifications typically examines the reliability of identifications made by bystanders (cf. Kassin, 1984). However, in the current studies, the eyewitness was the victim in both crimes (attempted rape and robbery). Jurors may have focused on the eyewitness as a victim of a violent crime where the perpetrator was up close and personal. Thus, regardless of the actual witnessing conditions, jurors may have just believed the eyewitness given the nature of the crimes. Future researchers should disentangle whether jurors perceive eyewitnesses differently depending on whether they are a bystander or victim. The role of sympathy should also be examined.

Furthermore, additional research should address whether victim eyewitnesses are more or less reliable than bystander eyewitnesses.
Other issues in the current studies should be teased apart. In Studies 1 and 2, I manipulated blocks of eyewitness factors in order to determine whether safeguards could draw jurors’ attention to identifications made under reliable or unreliable conditions. Future researchers should examine these factors to determine whether jurors are sensitive to each factor’s *individual* impact on accuracy and what, if any, additive effects there may be. Desmarais and Read’s (2009) survey of surveys provides general comprehension levels for various eyewitness factors, which can serve as a guide to test whether verdict decisions are differentially impacted by factors that are more or less understood by the majority of laypersons.

Additionally, researchers should continue testing a variety of safeguards, but borrow from other fields, such as linguistics (see Severance & Loftus, 1982) to shape the design of such safeguards. The results of the current studies indicate comprehension of both general (Study 1; cf., lineup instructions) and case-specific (Study 2) eyewitness factors is not improved with *Henderson* instructions (see also Bornstein & Hamm, 2012). Thus, particular focus should be paid in the future to improving jurors’ comprehension as has been done with pattern (Severance & Loftus, 1982) and death penalty instructions (Frank & Applegate, 1998).

Beyond judicial instructions, attorneys have a role in changing jurors’ perceptions of eyewitness evidence. As mentioned in Study 1, judges have discretion when determining which factors to include in the *Henderson* instructions, which may be difficult to accomplish given the subjective nature of many eyewitness factors (in particular, estimator variables). As with other instructions, attorneys may argue for the inclusion/exclusion of specific information in an instruction. Thus, (defense) attorneys must educate themselves on eyewitness factors if they wish to persuade the judge to include problematic factors. While it appears defense attorneys are more
adept at evaluating eyewitness evidence than prosecutors, there is room for improvement (Wise et al., 2009). Furthermore, defense attorneys have the opportunity to educate jurors during closing arguments and help draw attention to the quality of the eyewitness identification. However, evidence is mixed on whether closing arguments can accomplish these aims (Laub et al., 2014; Geiselman & Mendez, 2005). Future research should continue testing whether closing arguments can be an effective way to educate and sensitize jurors to eyewitness identification quality and whether attorneys who tie in the Henderson instructions with their arguments can induce a sensitivity effect.

As a leader in criminal law, the New Jersey Supreme Court’s decision to implement case-specific eyewitness instructions has already induced other states to follow suit (e.g., Utah, Connecticut, and Massachusetts). For example, the Massachusetts Supreme Court recently determined that jurors should receive eyewitness instructions on case-specific factors “for which there is at least a near consensus in the relevant scientific community” (Commonwealth v. Gomes, 2015, p. 376). The widespread adoption of instructions similar to Henderson should hinge on their effectiveness. Thus, given the current studies’ results, courts may wish to delay implementation of these instructions until further research can establish that these or some variant of these instructions actually improve jurors’ decisions above and beyond the decisions jurors would come to on their own.

While courts, such as New Jersey have been receptive to psychological research and allowed it to inform their decisions, it remains to be seen whether they will be flexible as our understanding of the impact of such instructions evolves. The number of wrongful convictions demonstrates the need for better, more effective safeguards to assist jurors in evaluating
eyewitness evidence. Chief Justice Rabner of the New Jersey Supreme Court stated himself, “To be effective, [instructions] cannot rely on a dated, analytical framework that has lost some of its vitality. Rather, they must be informed by sound evidence on memory and eyewitness identification, which is generally accepted by the relevant scientific community. Only then can courts fulfill their obligation both to defendants and the public” (New Jersey v. Henderson, 2011).
Footnotes

1. There were no gender differences across all outcome measures.
2. I also ran additional analyses with the control condition as the reference group. Results indicated differences only between the original *Henderson* instructions and both the control condition and modified *Henderson* instructions. The control and modified *Henderson* instructions did not render any differences in the logistic regression analyses.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Control</th>
<th>Henderson</th>
<th>Research</th>
<th>Expert</th>
<th>Expert + H</th>
<th>Poor</th>
<th>Good</th>
<th>Poor</th>
<th>Good</th>
<th>Participant Type</th>
<th>Community Member</th>
<th>Student</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension ratings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.01)</td>
<td>(1.80)</td>
<td>(1.98)</td>
<td>(1.78)</td>
<td>(1.48)</td>
<td></td>
<td>(1.69)</td>
<td>(1.90)</td>
<td>(1.82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No weapon</td>
<td>5.17</td>
<td>5.73</td>
<td>5.59</td>
<td>6.14*</td>
<td>5.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.70</td>
<td>5.67</td>
</tr>
<tr>
<td>(2.20)</td>
<td>(2.35)</td>
<td>(2.32)</td>
<td>(2.07)</td>
<td>(2.30)</td>
<td></td>
<td>(2.22)</td>
<td>(2.30)</td>
<td>(2.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short time delay</td>
<td>6.42</td>
<td>6.11</td>
<td>6.45</td>
<td>5.88</td>
<td>6.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.54</td>
<td>6.24</td>
</tr>
<tr>
<td>(2.51)</td>
<td>(2.40)</td>
<td>(2.41)</td>
<td>(2.55)</td>
<td>(2.61)</td>
<td></td>
<td>(2.38)</td>
<td>(2.57)</td>
<td>(2.49)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lineup (vs. showup)</td>
<td>5.56</td>
<td>6.27</td>
<td>5.70</td>
<td>5.79</td>
<td>5.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.08</td>
<td>5.82</td>
</tr>
<tr>
<td>(2.49)</td>
<td>(2.09)</td>
<td>(2.27)</td>
<td>(2.41)</td>
<td>(2.51)</td>
<td></td>
<td>(2.35)</td>
<td>(2.34)</td>
<td>(2.36)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unbiased instructions</td>
<td>4.62</td>
<td>5.78**</td>
<td>5.85**</td>
<td>5.84**</td>
<td>5.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.87</td>
<td>5.49</td>
</tr>
<tr>
<td>(2.37)</td>
<td>(2.57)</td>
<td>(2.33)</td>
<td>(2.18)</td>
<td>(2.39)</td>
<td></td>
<td>(2.37)</td>
<td>(2.39)</td>
<td>(2.41)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td>7.42</td>
<td>7.35</td>
<td>6.98</td>
<td>7.49</td>
<td>6.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.52</td>
<td>6.97**</td>
</tr>
<tr>
<td>(1.99)</td>
<td>(1.90)</td>
<td>(2.04)</td>
<td>(1.82)</td>
<td>(2.11)</td>
<td></td>
<td>(1.76)</td>
<td>(2.12)</td>
<td>(1.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructions help</td>
<td>6.96</td>
<td>7.38</td>
<td>7.01</td>
<td>7.07</td>
<td>7.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.16</td>
<td>7.15</td>
</tr>
<tr>
<td>(1.74)</td>
<td>(1.96)</td>
<td>(1.72)</td>
<td>(1.89)</td>
<td>(1.79)</td>
<td></td>
<td>(1.85)</td>
<td>(1.84)</td>
<td>(1.84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detective Cred</td>
<td>5.43</td>
<td>5.03</td>
<td>5.19</td>
<td>4.81**</td>
<td>5.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.07</td>
<td>5.12</td>
</tr>
<tr>
<td>(1.16)</td>
<td>(1.38)</td>
<td>(1.15)</td>
<td>(1.29)</td>
<td>(1.25)</td>
<td></td>
<td>(1.32)</td>
<td>(1.21)</td>
<td>(1.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyewitness Cred</td>
<td>5.19</td>
<td>5.02</td>
<td>5.04</td>
<td>4.86</td>
<td>5.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.20</td>
<td>4.95*</td>
</tr>
<tr>
<td>(1.40)</td>
<td>(1.29)</td>
<td>(1.35)</td>
<td>(1.15)</td>
<td>(1.27)</td>
<td></td>
<td>(1.29)</td>
<td>(1.29)</td>
<td>(1.29)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belief in eye</td>
<td>61.57</td>
<td>54.62</td>
<td>50.96</td>
<td>51.37</td>
<td>57.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>58.13</td>
<td>52.59*</td>
</tr>
<tr>
<td>(30.08)</td>
<td>(31.16)</td>
<td>(34.29)</td>
<td>(30.26)</td>
<td>(30.22)</td>
<td>(31.50)</td>
<td>(31.31)</td>
<td>(31.70)</td>
<td>(31.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verdict</td>
<td>.53 (.50)</td>
<td>.52 (.50)**</td>
<td>.46 (.50)**</td>
<td>.30 (.46)**</td>
<td>.51 (.50)</td>
<td>.44 (.50)</td>
<td>.48 (.50)</td>
<td>.41 (.49)</td>
<td>.52 (.50)*</td>
<td>.50 (.50)</td>
<td>.46 (.50)</td>
<td>.46 (.50)</td>
<td></td>
</tr>
</tbody>
</table>

*p = .05    **p = .01

Note. Verdict: 0=Not Guilty, 1 = Guilty. Comprehension ratings: 9 point scale, higher values = increased comprehension. Witness ratings: 7 point scale, higher values = increased credibility. Belief in eyewitness: 0-100, higher values = more likely a correct identification. Instructions help: 9 point scale, higher values = more agreement. Any significant differences within safeguard type are compared to control except for verdict where the proportion for expert is different than all other safeguard types, including control.
Table 2

*Study 1 Correlation matrix among proposed dependent and mediator variables (N = 448-452)*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Verdict</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comprehension ratings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Long duration</td>
<td>.16**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. No weapon</td>
<td>-.22**</td>
<td>-.14**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Short time delay</td>
<td>.02</td>
<td>.27**</td>
<td>-.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Lineup (vs. showup)</td>
<td>-.15**</td>
<td>-.13**</td>
<td>.21**</td>
<td>-.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Unbiased instructions</td>
<td>-.15**</td>
<td>-.09</td>
<td>.20**</td>
<td>-.01</td>
<td>.23**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. No feedback</td>
<td>-.09</td>
<td>.07</td>
<td>.13**</td>
<td>.15**</td>
<td>.08</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Witness ratings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Detective</td>
<td>.34**</td>
<td>.16**</td>
<td>-.12*</td>
<td>.10*</td>
<td>-.10*</td>
<td>-.04</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Eyewitness</td>
<td>.54**</td>
<td>.22**</td>
<td>-.14**</td>
<td>.12**</td>
<td>-.13**</td>
<td>-.08</td>
<td>.07</td>
<td>.59**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Belief in eye</td>
<td>.74**</td>
<td>.26**</td>
<td>-.22**</td>
<td>.10*</td>
<td>-.25**</td>
<td>-.15%</td>
<td>.01</td>
<td>.37**</td>
<td>.60**</td>
<td></td>
</tr>
<tr>
<td>10. Instructions help</td>
<td>.09</td>
<td>.16**</td>
<td>-.05</td>
<td>.13**</td>
<td>-.04</td>
<td>.05</td>
<td>.08</td>
<td>.16**</td>
<td>.18**</td>
<td>.16**</td>
</tr>
</tbody>
</table>

*Note. Verdict: 0=not guilty; 1 = guilty. Comprehension ratings: 9 point scale with higher values = increased comprehension. Witness ratings: 7 point scale with higher values = increased credibility. Belief in eyewitness: 0-100 with higher values = more likely a correct identification. Instructions help: 9 point scale, higher values = more agreement.*

*p = .05

**p = .01
Table 3

Study 1 Sensitivity tests: Dependent variables as a function of safeguard type, and witnessing and identification conditions (N=448-452)

<table>
<thead>
<tr>
<th>Safeguard Type</th>
<th>Dependent Variable</th>
<th>Control</th>
<th>Henderson Research</th>
<th>Expert</th>
<th>Expert + H</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M ) (SD)</td>
<td>( M ) (SD)</td>
<td>( M ) (SD)</td>
<td>( M ) (SD)</td>
<td>( M ) (SD)</td>
<td>( M ) (SD)</td>
</tr>
<tr>
<td><strong>Detective ratings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good W</td>
<td>5.20 (1.24)</td>
<td>5.10 (1.50)</td>
<td>5.17 (1.10)</td>
<td>4.83 (1.37)</td>
<td>5.00 (1.22)</td>
</tr>
<tr>
<td>Poor W</td>
<td>5.63 (1.06)</td>
<td>4.94 (1.25)</td>
<td>5.20 (1.21)</td>
<td>4.79 (1.23)</td>
<td>5.11 (1.30)</td>
</tr>
<tr>
<td>Good ID</td>
<td>5.60 (1.20)</td>
<td>5.49 (1.30)</td>
<td>5.19 (1.21)</td>
<td>5.24 (1.13)</td>
<td>5.37 (1.19)</td>
</tr>
<tr>
<td>Poor ID</td>
<td>5.25 (1.10)</td>
<td>4.57 (1.31)</td>
<td>5.18 (1.10)</td>
<td>4.45 (1.32)</td>
<td>4.72 (1.23)</td>
</tr>
<tr>
<td>Good W, Good ID</td>
<td>5.45 (1.28)</td>
<td>5.63 (1.41)</td>
<td>5.31 (1.00)</td>
<td>5.43 (1.16)</td>
<td>5.07 (1.22)</td>
</tr>
<tr>
<td>Good W, Poor ID</td>
<td>4.84 (1.11)</td>
<td>4.63 (1.45)</td>
<td>4.99 (1.21)</td>
<td>4.39 (1.36)</td>
<td>4.94 (1.24)</td>
</tr>
<tr>
<td>Poor W, Good ID</td>
<td>5.75 (1.12)</td>
<td>5.35 (2.21)</td>
<td>5.07 (1.39)</td>
<td>5.08 (1.10)</td>
<td>5.71 (1.09)</td>
</tr>
<tr>
<td>Poor W, Poor ID</td>
<td>5.53 (1.02)</td>
<td>4.51 (1.16)</td>
<td>5.35 (1.00)</td>
<td>4.51 (1.30)</td>
<td>4.45 (1.21)</td>
</tr>
<tr>
<td><strong>Eyewitness ratings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good W</td>
<td>5.14 (1.52)</td>
<td>5.21 (1.23)</td>
<td>5.24 (1.24)</td>
<td>5.07 (1.09)</td>
<td>5.20 (1.13)</td>
</tr>
<tr>
<td>Poor W</td>
<td>5.23 (1.30)</td>
<td>4.82 (1.33)</td>
<td>4.86 (1.44)</td>
<td>4.66 (1.17)</td>
<td>5.26 (1.41)</td>
</tr>
<tr>
<td>Good ID</td>
<td>5.06 (1.33)</td>
<td>5.40 (1.34)</td>
<td>5.18 (1.42)</td>
<td>5.04 (1.18)</td>
<td>5.37 (1.39)</td>
</tr>
<tr>
<td>Poor ID</td>
<td>5.33 (1.47)</td>
<td>4.65 (1.13)</td>
<td>4.88 (1.27)</td>
<td>4.72 (1.11)</td>
<td>5.08 (1.12)</td>
</tr>
<tr>
<td>Good W, Good ID</td>
<td>5.06 (1.46)</td>
<td>5.65 (1.15)</td>
<td>5.46 (1.19)</td>
<td>5.37 (1.09)</td>
<td>5.07 (1.38)</td>
</tr>
<tr>
<td>Good W, Poor ID</td>
<td>5.26 (1.64)</td>
<td>4.80 (1.18)</td>
<td>4.96 (1.28)</td>
<td>4.84 (1.06)</td>
<td>5.32 (0.83)</td>
</tr>
<tr>
<td>Poor W, Good ID</td>
<td>5.07 (1.22)</td>
<td>5.14 (1.49)</td>
<td>4.90 (1.59)</td>
<td>4.75 (1.19)</td>
<td>5.70 (1.36)</td>
</tr>
<tr>
<td>Poor W, Poor ID</td>
<td>5.38 (1.38)</td>
<td>4.49 (1.07)</td>
<td>4.82 (1.28)</td>
<td>4.58 (1.17)</td>
<td>4.80 (1.35)</td>
</tr>
<tr>
<td><strong>Belief in eyewitness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good W</td>
<td>61.90 (30.64)</td>
<td>55.49 (31.85)</td>
<td>55.02 (32.66)</td>
<td>49.06 (31.01)</td>
<td>52.50 (30.33)</td>
</tr>
<tr>
<td>Poor W</td>
<td>61.29 (29.91)</td>
<td>53.70 (30.76)</td>
<td>47.23 (35.66)</td>
<td>53.68 (29.64)</td>
<td>63.46 (29.39)</td>
</tr>
<tr>
<td>Good ID</td>
<td>61.45 (28.32)</td>
<td>58.32 (32.35)</td>
<td>52.72 (34.41)</td>
<td>51.86 (29.98)</td>
<td>58.93 (29.73)</td>
</tr>
<tr>
<td>Poor ID</td>
<td>61.71 (32.28)</td>
<td>51.01 (29.88)</td>
<td>48.86 (34.44)</td>
<td>50.96 (30.78)</td>
<td>56.37 (31.01)</td>
</tr>
<tr>
<td>Good W, Good ID</td>
<td>62.67 (31.88)</td>
<td>57.36 (33.38)</td>
<td>55.44 (31.72)</td>
<td>54.00 (29.69)</td>
<td>54.57 (30.89)</td>
</tr>
<tr>
<td>Good W, Poor ID</td>
<td>60.82 (29.74)</td>
<td>53.77 (31.00)</td>
<td>54.47 (34.72)</td>
<td>52.07 (32.17)</td>
<td>50.43 (30.31)</td>
</tr>
<tr>
<td>Poor W, Good ID</td>
<td>60.17 (24.72)</td>
<td>59.27 (32.04)</td>
<td>50.00 (37.37)</td>
<td>57.83 (29.57)</td>
<td>63.71 (28.38)</td>
</tr>
<tr>
<td>Poor W, Poor ID</td>
<td>62.32 (34.49)</td>
<td>47.86 (28.96)</td>
<td>44.22 (34.28)</td>
<td>49.71 (29.77)</td>
<td>63.20 (31.15)</td>
</tr>
<tr>
<td><strong>Verdict</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good W</td>
<td>.51 (.51)</td>
<td>.60 (.50)</td>
<td>.56 (.50)</td>
<td>.28 (.46)</td>
<td>.47 (.51)</td>
</tr>
<tr>
<td>Poor W</td>
<td>.54 (.50)</td>
<td>.44 (.50)</td>
<td>.37 (.49)</td>
<td>.32 (.47)</td>
<td>.55 (.50)</td>
</tr>
<tr>
<td>Good ID</td>
<td>.53 (.50)</td>
<td>.60 (.50)</td>
<td>.48 (.51)</td>
<td>.40 (.50)</td>
<td>.58 (.50)</td>
</tr>
<tr>
<td>Poor ID</td>
<td>.52 (.51)</td>
<td>.44 (.50)</td>
<td>.44 (.50)</td>
<td>.22 (.42)</td>
<td>.43 (.50)</td>
</tr>
<tr>
<td>Good W, Good ID</td>
<td>.54 (.51)</td>
<td>.71 (.46)</td>
<td>.56 (.51)</td>
<td>.40 (.50)</td>
<td>.55 (.51)</td>
</tr>
<tr>
<td>Good W, Poor ID</td>
<td>.47 (.51)</td>
<td>.50 (.51)</td>
<td>.55 (.51)</td>
<td>.19 (.40)</td>
<td>.39 (.50)</td>
</tr>
<tr>
<td>Poor W, Good ID</td>
<td>.52 (.51)</td>
<td>.50 (.51)</td>
<td>.40 (.50)</td>
<td>.39 (.50)</td>
<td>.62 (.50)</td>
</tr>
<tr>
<td>Poor W, Poor ID</td>
<td>.56 (.51)</td>
<td>.38 (.50)</td>
<td>.35 (.49)</td>
<td>.25 (.44)</td>
<td>.47 (.51)</td>
</tr>
</tbody>
</table>

*Note: W=watching conditions; ID=identification conditions. Good/Poor W is collapsed across ID conditions. Good/Poor ID is collapsed across witnessing conditions.*
<table>
<thead>
<tr>
<th>Variable</th>
<th>Total</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$, (SE)</td>
<td>$b$, (SE)</td>
<td>$b$, (SE)</td>
</tr>
<tr>
<td></td>
<td>[CI]</td>
<td>[CI]</td>
<td>[CI]</td>
</tr>
<tr>
<td>Quality of ID conditions</td>
<td>.46, (.19)*</td>
<td>.21, (.24)</td>
<td>.32, (.14)*</td>
</tr>
<tr>
<td></td>
<td>[.07, .83]</td>
<td>[-.26, .67]</td>
<td>[.04, .61]</td>
</tr>
<tr>
<td>Detective credibility</td>
<td>.05, (.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-.17, .28]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyewitness credibility</td>
<td>1.09, (.13)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[.84, 1.34]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-2LL = 597.24           -2LL = 455.71
Nagelkerke $R^2 = .06$   Nagelkerke $R^2 = .40$

*p < .05.

Note. The significant indirect effect indicated that the effect of quality of identification conditions on verdict was mediated by eyewitness credibility. Quality of witnessing conditions and the four dummy coded variables comprising safeguard type were included as covariates.
### Study 1 Mediation model with expert testimony predicting conviction (PROCESS Model 4)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$, (SE)</td>
<td>$b$, (SE)</td>
<td>$b$, (SE)</td>
</tr>
<tr>
<td></td>
<td>[CI]</td>
<td>[CI]</td>
<td>[CI]</td>
</tr>
<tr>
<td>Expert only(^1)</td>
<td>-.94, (.31)*</td>
<td>-.71, (.33)*</td>
<td>-.34, (.12)*</td>
</tr>
<tr>
<td></td>
<td>[-1.56, -.33]</td>
<td>[-1.36, -.07]</td>
<td>[-.60, -.13]</td>
</tr>
<tr>
<td>Detective credibility</td>
<td>.58, (.09)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[.41, .76]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-2LL = 597.24
Nagelkerke $R^2 = .06$

-2LL = 550.12
Nagelkerke $R^2 = .19$

\(^*p < .05.\)

**Note.** The significant indirect effect indicated that the effect of expert testimony (vs. no safeguard) on verdict was mediated by detective credibility.

\(^1\) Expert only was one of four dummy coded variables comprising safeguard type. No safeguard was the reference group. The remaining dummy coded safeguard variables, as well as quality of witnessing and identification conditions were included as covariates.
Table 6  
Study 2 Descriptives (N=474-482)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Instruction Type M (SD)</th>
<th>Evidence Strength M (SD)</th>
<th>Interrogatories Timing M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Original</td>
<td>Modified</td>
</tr>
<tr>
<td>Comprehension</td>
<td>5.42</td>
<td>5.06</td>
<td>5.26</td>
</tr>
<tr>
<td></td>
<td>(1.98)</td>
<td>(1.91)</td>
<td>(2.11)</td>
</tr>
<tr>
<td>Instructions help</td>
<td>7.08</td>
<td>7.82**</td>
<td>7.75**</td>
</tr>
<tr>
<td></td>
<td>(1.79)</td>
<td>(1.37)</td>
<td>(1.31)</td>
</tr>
<tr>
<td>Witness ratings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detective</td>
<td>5.51</td>
<td>5.70</td>
<td>5.70</td>
</tr>
<tr>
<td></td>
<td>(1.28)</td>
<td>(1.07)</td>
<td>(1.17)</td>
</tr>
<tr>
<td>Eyewitness</td>
<td>5.50</td>
<td>5.39</td>
<td>5.51</td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(1.25)</td>
<td>(1.44)</td>
</tr>
<tr>
<td>Belief in eye</td>
<td>68.41</td>
<td>61.41</td>
<td>66.01</td>
</tr>
<tr>
<td></td>
<td>(26.64)</td>
<td>(25.07)</td>
<td>(27.13)</td>
</tr>
<tr>
<td>Verdict</td>
<td>.58 (.50)</td>
<td>.41 (.49)</td>
<td>.52 (.50)</td>
</tr>
</tbody>
</table>

* p = .01  **p = .05

Note. Comprehension ratings: 9 point scale with higher values = increased comprehension of the impact of three estimator variables on accuracy. Witness ratings: 7 point scale with higher values = increased credibility. Belief in eyewitness: 0-100 with higher values = increased likelihood of correct identification. Verdict: 0=Not Guilty, 1 = Guilty. Instructions help: 9 point scale, higher values = more agreement. Any significant differences within instruction type are compared to control.
Table 7
Study 2 Correlation matrix among proposed mediators and dependent variables (N=474-482)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Verdict</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comprehension ratings</td>
<td>.56*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Detective ratings</td>
<td>.34*</td>
<td>.30*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Eyewitness ratings</td>
<td>.53*</td>
<td>.44*</td>
<td>.67*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Belief in eyewitness</td>
<td>.74*</td>
<td>.65*</td>
<td>.44*</td>
<td>.63*</td>
<td></td>
</tr>
<tr>
<td>5. Instructions help</td>
<td>-.00</td>
<td>.05</td>
<td>.21*</td>
<td>.12*</td>
<td>.05</td>
</tr>
</tbody>
</table>

*p = .01

Note. Comprehension ratings: 9 point scale with higher values = increased comprehension of the impact of three estimator variables on accuracy. Witness ratings: 7 point scale with higher values = increased credibility. Belief in eyewitness: 0-100 with higher values = increased likelihood of correct identification. Verdict: 0=Not Guilty, 1 = Guilty. Instructions help: 9 point scale, higher values = more agreement.
Table 8
Study 2 Sensitivity tests: Dependent variables as a function of instruction type and evidence strength (N=474-482)

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Control</th>
<th>Original</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comprehension ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak evidence</td>
<td>4.11 (1.77)</td>
<td>3.83 (1.76)</td>
<td>3.71 (1.66)</td>
</tr>
<tr>
<td>Strong evidence</td>
<td>6.52 (1.41)</td>
<td>6.00 (1.44)</td>
<td>6.52 (1.51)</td>
</tr>
<tr>
<td><strong>Detective ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak evidence</td>
<td>5.38 (1.27)</td>
<td>5.82 (1.11)</td>
<td>5.68 (1.15)</td>
</tr>
<tr>
<td>Strong evidence</td>
<td>5.61 (1.29)</td>
<td>5.60 (1.04)</td>
<td>5.72 (1.19)</td>
</tr>
<tr>
<td><strong>Eyewitness ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak evidence</td>
<td>5.26 (1.44)</td>
<td>5.37 (1.53)</td>
<td>5.19 (1.56)</td>
</tr>
<tr>
<td>Strong evidence</td>
<td>5.71 (1.46)</td>
<td>5.41 (1.00)</td>
<td>5.77 (1.27)</td>
</tr>
<tr>
<td><strong>Belief in eyewitness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak evidence</td>
<td>59.57 (26.54) &lt;sub&gt;b&lt;/sub&gt;</td>
<td>61.37 (25.83)</td>
<td>55.15 (28.34) &lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Strong evidence</td>
<td>75.83 (24.49) &lt;sub&gt;b&lt;/sub&gt;</td>
<td>61.44 (24.72)</td>
<td>74.86 (22.65) &lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td><strong>Instructions help</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak evidence</td>
<td>7.29 (1.79)</td>
<td>7.38 (1.64) &lt;sub&gt;a&lt;/sub&gt;</td>
<td>7.74 (1.32)</td>
</tr>
<tr>
<td>Strong evidence</td>
<td>6.91 (1.78)</td>
<td>8.15 (1.01) &lt;sub&gt;a&lt;/sub&gt;</td>
<td>7.96 (1.13)</td>
</tr>
<tr>
<td><strong>Verdict</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak evidence</td>
<td>.46 (.50) &lt;sub&gt;b&lt;/sub&gt;</td>
<td>.45 (.50)</td>
<td>.35 (.48) &lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Strong evidence</td>
<td>.68 (.47) &lt;sub&gt;b&lt;/sub&gt;</td>
<td>.38 (.49)</td>
<td>.66 (.48) &lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

*Note: Within dependent measures, means sharing subscripts differ at p < .05.*
Figure 1. Mediational model displaying the effect of system variables on verdict mediated by eyewitness credibility ratings (PROCESS Model 4; Hayes, 2013).

Total effect, $b = .45$, 95% CI $[.07, .83]$  
Direct effect, $b = .21$, 95% CI $[-.26, .67]$  
Indirect effect, $b = .32$, 95% CI $[.04, .61]$  

$Eyewitness Credibility$  

$System Variables$  

$Verdict$  

$b = .29$, 95% CI $[.05, .53]$  

$b = 1.09$, 95% CI $[.84, 1.34]$
Figure 2. Mediational model displaying the effect of expert testimony on verdict mediated by detective credibility ratings (PROCESS Model 4; Hayes, 2013).

Total effect, $b = -0.94$, 95% CI [-1.56, -.33]
Direct effect, $b = -0.71$, 95% CI [-1.36, -.07]
Indirect effect, $b = -0.34$, 95% CI [-0.60, -0.13]

$\hat{b} = -0.58$, 95% CI [-0.93, -0.22]
$\hat{b} = 0.58$, 95% CI [0.41, 0.76]
Figure 3. Moderated Mediational model displaying the conditional effect of evidence strength on verdict mediated by belief in the eyewitness (PROCESS Model 10; Hayes, 2013).

Direct effect, $b = -0.15$, 95% CI [-0.73, 0.43]  
Conditional indirect effect within control condition, $b = 1.74$, 95% CI [0.93, 2.60]  
Conditional indirect effect within modified instructions condition, $b = 2.11$, 95% CI [1.27, 3.08]
Appendix A: Original *Henderson* Instructions with Research Enhancement in Brackets

**IDENTIFICATION: IN-COURT AND OUT-OF-COURT IDENTIFICATIONS**

[Defendant] as part of his general denial of guilt, contends that the State has not presented sufficient reliable evidence to establish beyond a reasonable doubt that he is the person who committed the alleged offense. The defendant has neither the burden nor the duty to show that the crime, if committed, was committed by someone else, or to prove the identity of that other person. You must determine, therefore, not only whether the State has proven each and every element of the offense charged beyond a reasonable doubt, but also whether the State has proven beyond a reasonable doubt that this defendant is the person who committed it.

The State has presented the testimony of [eyewitness]. You will recall that this witness identified the defendant in court as the person who committed [charge(s)]. The State also presented testimony that on a prior occasion before this trial, this witness identified the defendant as the person who committed these offenses. According to the witness, [his/her] identification of the defendant was based upon the observations and perceptions that s/he made of the perpetrator at the time the offense was being committed. It is your function to determine whether the witness’s identification of the defendant is reliable and believable, or whether it is based on a mistake or for any reason is not worthy of belief. You must decide whether it is sufficiently reliable evidence that this defendant is the person who committed the offense(s) charged.

Eyewitness identification evidence must be scrutinized carefully. Human beings have the ability to recognize other people from past experiences and to identify them at a later time, but research has shown that there are risks of making mistaken identifications. That research has focused on the nature of memory and the factors that affect the reliability of eyewitness identifications.

Human memory is not foolproof. Research has revealed that human memory is not like a video recording that a witness need only replay to remember what happened. Memory is far more complex. The process of remembering consists of three stages: acquisition -- the perception of the original event; retention -- the period of time that passes between the event and the eventual recollection of a piece of information; and retrieval -- the stage during which a person recalls stored information. At each of these stages, memory can be affected by a variety of factors.

Relying on some of the research that has been done, I will instruct you on specific factors you should consider in this case in determining whether the eyewitness identification evidence is reliable. In evaluating this identification, you should consider the observations and perceptions on which the identification was based, the witness’s ability to make those observations and perceive events, and the circumstances under which the
identification was made. Although nothing may appear more convincing than a witness’s categorical identification of a perpetrator, you must critically analyze such testimony. Such identifications, even if made in good faith, may be mistaken. Therefore, when analyzing such testimony, be advised that a witness’s level of confidence, standing alone, may not be an indication of the reliability of the identification.

If you determine that the out-of-court identification is not reliable, you may still consider the witness’s in-court identification of the defendant if you find that it resulted from the witness’s observations or perceptions of the perpetrator during the commission of the offense, and that the identification is reliable. If you find that the in-court identification is the product of an impression gained at the out-of-court identification procedure, it should be afforded no weight. The ultimate question of the reliability of both the in-court and out-of-court identifications is for you to decide.

To decide whether the identification testimony is sufficiently reliable evidence to conclude that this defendant is the person who committed the offense(s) charged, you should evaluate the testimony of the witness in light of the factors for considering credibility that I have already explained to you. In addition, you should consider the following factors that are related to the witness, the alleged perpetrator, and the criminal incident itself. In particular, you should consider:

(1) The Witness’s Opportunity to View and Degree of Attention: In evaluating the reliability of the identification, you should assess the witness’s opportunity to view the person who committed the offense at the time of the offense and the witness’s degree of attention to the perpetrator at the time of the offense. In making this assessment you should consider the following:

**Stress:** Even under the best viewing conditions, high levels of stress can reduce an eyewitness’s ability to recall and make an accurate identification. Therefore, you should consider a witness’s level of stress and whether that stress, if any, distracted the witness or made it harder for him or her to identify the perpetrator. [Research enhanced version: Additionally a recent meta-analysis demonstrates that when the perpetrator is present in the lineup, the identification accuracy rate under high stress was only 39% compared to 59% under low stress.]

**Duration:** The amount of time an eyewitness has to observe an event may affect the reliability of an identification. Although there is no minimum time required to make an accurate identification, a brief or fleeting contact is less likely to produce an accurate identification than a more prolonged exposure to the perpetrator. In addition, time estimates given by witnesses may not always be accurate because witnesses tend to think events lasted longer than they actually did. [Research enhanced version: Research demonstrates reliable differences between identifications made when the witness sees the perpetrator for 12 seconds and 45 seconds. Those viewing a perpetrator for only 12 seconds made correct identifications in 30% of perpetrator present lineups and false identifications in 85% of perpetrator absent lineups. On the other hand, those viewing a perpetrator for 45 seconds had 90% correct identifications in perpetrator present lineups and 45% false identifications in perpetrator absent lineups.]
**Weapon Focus:** You should consider whether the witness saw a weapon during the incident and the duration of the crime. The presence of a weapon can distract the witness and take the witness’s attention away from the perpetrator's face. As a result, the presence of a visible weapon may reduce the reliability of a subsequent identification if the crime is of short duration. In considering this factor, you should take into account the duration of the crime because the longer the event, the more time the witness may have to adapt to the presence of the weapon and focus on other details. [Research enhanced version: One study found that the presence of a weapon increased false identifications from 33% to 64%.]

**Lighting:** Inadequate lighting can reduce the reliability of an identification. You should consider the lighting conditions present at the time of the alleged crime in this case. [Research enhanced version: Research demonstrates that identification accuracy tend to be approximately 50% in poor lighting conditions, compared to almost 90% in well-lit conditions.]

**Disguises/Changed Appearance:** The perpetrator’s use of a disguise can affect a witness’s ability both to remember and identify the perpetrator. Disguises like hats, sunglasses, or masks can reduce the accuracy of an identification. Similarly, if facial features are altered between the time of the event and a later identification procedure, the accuracy of the identification may decrease. [Research enhanced version: One study found that a perpetrator of a mock crime wearing a hat was identified by only 27% of participants, compared to 45% of participants when the perpetrator was not wearing a hat.]

**Prior Description of Perpetrator:** Another factor for your consideration is the accuracy of any description the witness gave after observing the incident and before identifying the perpetrator. Facts that may be relevant to this factor include whether the prior description matched the photo or person picked out later, whether the prior description provided details or was just general in nature, and whether the witness's testimony at trial was consistent with, or different from, her prior description of the perpetrator. You may also consider whether the witness did not identify the defendant at a prior identification procedure or chose a different suspect or filler.

**Confidence and Accuracy:** You heard testimony that Marie Sands made a statement at the time she identified the defendant concerning her level of certainty that the person she selected is in fact the person who committed the crime. As I explained earlier, a witness’s level of confidence, standing alone, may not be an indication of the reliability of the identification. Although some research has found that highly confident witnesses are more likely to make accurate identifications, eyewitness confidence is generally an unreliable indicator of accuracy. [Research enhanced version: It is not unusual for witnesses to make confidence judgments that are 15-20% higher than actual accuracy rates.]

**Time Elapsed:** Memories fade with time. As a result, delays between the commission of a crime and the time an identification is made can affect the reliability of the identification. In other words, the more time that passes, the greater the possibility that a witness’s memory of a perpetrator will weaken. If there is evidence of impairment by drugs or other substances, the
charge can be modified accordingly on. [Research enhanced version: A review of the literature found that longer delays can lead to 10% fewer correct identifications and 8% more false identifications. Specifically, research has demonstrated that eyewitness accuracy declines 20-25% in the first 2 hours. In the next 10 hours, another 5% of memory loss is experienced with more gradual losses as time extends beyond that.]

**Cross-Racial Effects:** Research has shown that people may have greater difficulty in accurately identifying members of a different race. You should consider whether the fact that the witness and the defendant are not of the same race may have influenced the accuracy of the witness’s identification. [Research enhanced version: In particular, research indicates that overall accuracy rates are 60% for same-race identifications, compared to 40% for cross-race identifications.]

In evaluating the reliability of a witness’s identification, you should also consider the circumstances under which any out-of-court identification was made, and whether it was the result of a suggestive procedure. In that regard, you may consider everything that was done or said by law enforcement to the witness during the identification process. You should consider the following factors:

**Multiple Viewings:** When a witness views the same person in more than one identification procedure, it can be difficult to know whether a later identification comes from the witness’s memory of the actual, original event or of an earlier identification procedure. As a result, if a witness views an innocent suspect in multiple identification procedures, the risk of mistaken identification is increased. [Research enhanced version: A meta-analysis on mug shot effects found that seeing someone in a mug shot can double the risk of a subsequent misidentification. Specifically, false identifications were 16% for a previously unseen face, but went up to 38% when witnesses previously viewed a mug shot of that person.]

**Showups:** In this case, the witness identified the defendant during a “showup,” that is, the defendant was the only person shown to the witness at that time. Even though such a procedure is suggestive in nature, it is sometimes necessary for the police to conduct a “showup” or one-on-one identification procedure. Although the benefits of a fresh memory may balance the risk of undue suggestion, showups conducted more than two hours after an event present a heightened risk of misidentification. Also, police officers must instruct witnesses that the person they are about to view may or may not be the person who committed the crime and that they should not feel compelled to make an identification. In determining whether the identification is reliable or the result of an unduly suggestive procedure, you should consider how much time elapsed after the witness last saw the perpetrator, whether the appropriate instructions were given to the witness, and all other circumstances surrounding the showup. [Research enhanced version: A recent review of the literature found that false identifications are 23% in single person showups but only 10% in six person lineups.]

In determining the reliability of the identification, you should also consider whether the identification procedure was properly conducted.
Double-blind: A lineup administrator who knows which person or photo in the lineup is the suspect may intentionally or unintentionally convey that knowledge to the witness. That increases the chance that the witness will identify the suspect, even if the suspect is innocent. For that reason, whenever feasible, live lineups and photo arrays should be conducted by an officer who does not know the identity of the suspect.

If a police officer who does not know the suspect’s identity is not available, then the officer should not see the photos as the witness looks at them. In this case, it is alleged that the person who presented the lineup knew the identity of the suspect.

You may consider this factor when you consider the circumstances under which the identification was made, and when you evaluate the overall reliability of the identification.

Instructions: You should consider what was or what was not said to the witness prior to viewing a lineup or showup. Identification procedures should begin with instructions to the witness that the perpetrator may or may not be present and that the witness should not feel compelled to make an identification. The failure to give this instruction can increase the risk of misidentification. If you find that the police did or did not give this instruction to the witness, you may take this factor into account when evaluating the identification evidence. [Research enhanced version: A recent review of the literature found that when the witness was told that the perpetrator may or may not be present in lineups, error rates dropped from 60 to 35%.

Feedback: Feedback occurs when police officers, or witnesses to an event who are not law enforcement officials, signal to eyewitnesses that they correctly identified the suspect. That confirmation may reduce doubt and engender or produce a false sense of confidence in a witness. Feedback may also falsely enhance a witness’s recollection of the quality of his or her view of an event. It is for you to determine whether or not a witness’s recollection in this case was affected by feedback or whether the recollection instead reflects the witness’s accurate perception of the event. [Research enhanced version: A recent literature review demonstrates that confirmatory feedback about the accuracy of an identification can substantially increase confidence and inflate a witness’ sense of the quality of their memory.]

You may consider whether the witness was exposed to opinions, descriptions, or identifications given by other witnesses, to photographs or newspaper accounts, or to any other information or influence, that may have affected the independence of his/her identification. Such information can affect the independent nature and reliability of a witness’s identification and inflate the witness’s confidence in the identification. You are also free to consider any other factor based on the evidence or lack of evidence in the case that you consider relevant to your determination whether the identifications were reliable. Keep in mind that the presence of any single factor or combination of factors, however, is not an indication that a particular witness is incorrect. Instead, you may consider the factors that I have discussed as you assess all of the circumstances of the case, including all of the testimony and documentary evidence, in determining whether a particular identification made by a witness is accurate and thus worthy of your consideration as you decide whether the State has met its burden to prove identification beyond a reasonable doubt. If you determine that the in-court or out-of-court identifications resulted from the
witness's observations or perceptions of the perpetrator during the commission of the offense, you may consider that evidence and decide how much weight to give it. If you instead decide that the identification is or identifications are the product of an impression gained at the in-court and/or out-of-court identification procedures, the identifications should be afforded no weight. The ultimate issue of the trustworthiness of an identification is for you to decide.

If, after consideration of all of the evidence, you determine that the State has not proven beyond a reasonable doubt that [defendant] was the person who committed [this/these] offense(s), then you must find [her/him] not guilty. If, on the other hand, after consideration of all of the evidence, you are convinced beyond a reasonable doubt that [defendant] was correctly identified, you will then consider whether the State has proven each and every element of the offense(s) charged beyond a reasonable doubt.
Appendix B: Lay Eyewitness Knowledge (Desmarais & Read, 2011)

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>( r^2 )</th>
<th>( R^2 )</th>
<th>( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>25</td>
<td>0.90</td>
<td>0.84</td>
<td>0.90</td>
</tr>
<tr>
<td>21</td>
<td>26</td>
<td>0.92</td>
<td>0.85</td>
<td>0.92</td>
</tr>
<tr>
<td>22</td>
<td>27</td>
<td>0.93</td>
<td>0.86</td>
<td>0.93</td>
</tr>
<tr>
<td>23</td>
<td>28</td>
<td>0.94</td>
<td>0.87</td>
<td>0.94</td>
</tr>
<tr>
<td>24</td>
<td>29</td>
<td>0.95</td>
<td>0.88</td>
<td>0.95</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>0.96</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>26</td>
<td>31</td>
<td>0.97</td>
<td>0.90</td>
<td>0.97</td>
</tr>
<tr>
<td>27</td>
<td>32</td>
<td>0.98</td>
<td>0.91</td>
<td>0.98</td>
</tr>
<tr>
<td>28</td>
<td>33</td>
<td>0.99</td>
<td>0.92</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 1: Mean recall of percent correct answers with Klassen et al. (2010) excerpts.
Appendix C: Modified *Henderson* Instructions

The Court instructs you as follows:

  Eyewitness identification evidence must be scrutinized carefully. Research has shown that there are risks of making mistaken identifications. Human memory is not foolproof. Research has revealed that human memory is not like a video recording that a witness need only replay to remember what happened. Memory is far more complex.

  To decide whether the identification testimony is sufficiently reliable evidence to conclude that this defendant is the person who committed the offense charged, you should consider the following factors that are related to the witness, the alleged perpetrator, and the criminal incident itself. In particular, you should consider:

  (1) The Witness’s Opportunity to View and Degree of Attention: In evaluating the reliability of the identification, you should assess the witness’s opportunity to view the person who committed the offense at the time of the offense and the witness’s degree of attention to the perpetrator at the time of the offense. In making this assessment you should consider the following:

      (a) Stress: Even under the best viewing conditions, high levels of stress can reduce an eyewitness’s ability to recall and make an accurate identification. Therefore, you should consider a witness’s level of stress and whether that stress, if any, distracted the witness and made it harder for him or her to identify the perpetrator.

  **Ask Yourself: Was the eyewitness stressed? Does this enhance or impair his/her accuracy?**
(b) **Duration**: The amount of time an eyewitness has to observe an event may affect the reliability of an identification. Although there is no minimum time required to make an accurate identification, a brief or fleeting contact is less likely to produce an accurate identification than a more prolonged exposure to the perpetrator. In addition, time estimates given by witnesses may not always be accurate because witnesses tend to think events lasted longer than they actually did.

**Ask Yourself**: For how long did the eyewitness view the perpetrator? Does this enhance or impair his/her accuracy?

(c) **Weapon Focus**: You should consider whether the witness saw a weapon during the incident and the duration of the crime. The presence of a weapon can distract the witness’s attention away from the perpetrator's face. As a result, the presence of a visible weapon may reduce the reliability of a subsequent identification. In considering this factor, you should take into account the duration of the crime because the longer the event, the more time the witness may have to adapt to the presence of the weapon and focus on other details.

**Ask Yourself**: Did the perpetrator have a weapon? Does this enhance or impair his/her accuracy?

(d) **Distance**: A person is easier to identify when close by. The greater the distance between an eyewitness and a perpetrator, the higher the risk of a mistaken identification. In addition, a witness’s estimate of how far he or she was from the perpetrator may not always be accurate because people tend to have difficulty estimating distances.
Ask Yourself: How far apart were the perpetrator and the eyewitness? Does this enhance or impair his/her accuracy?

(e) Lighting: Inadequate lighting can reduce the reliability of an identification. You should consider the lighting conditions present at the time of the alleged crime in this case.

Ask Yourself: What were the lighting conditions during the crime? Does this enhance or impair his/her accuracy?

(f) Intoxication: The influence of alcohol can affect the reliability of an identification. An identification made by a witness under the influence of a high level of alcohol at the time of the incident tends to be less reliable than an identification by a witness who drank a small amount of alcohol or none at all.

Ask Yourself: Was the eyewitness intoxicated? Does this enhance or impair his/her accuracy?

(g) Disguises/Changed Appearance: The perpetrator’s use of a disguise can affect a witness’s ability both to remember and identify the perpetrator. Disguises like hats, sunglasses, or masks can reduce the accuracy of an identification. Similarly, if facial features are altered between the time of the event and a later identification procedure, the accuracy of the identification may decrease.

Ask Yourself: Was the perpetrator wearing a disguise or did their appearance change? Does this enhance or impair his/her accuracy?

(2) Prior Description of Perpetrator: Consider the accuracy of any description the witness gave after observing the incident and before identifying the perpetrator. Facts that
may be relevant to this factor include whether the prior description matched the photo or person picked out later, whether the prior description provided details or was just general in nature, and whether the witness's testimony at trial was consistent with, or different from, his/her prior description of the perpetrator.

**Ask Yourself:** Did the eyewitness’ prior description of the perpetrator match the defendant? Does this suggest the identification is more or less accurate?

**(3) Time Elapsed:** Memories fade with time. As a result, delays between the commission of a crime and the time an identification is made can affect the reliability of the identification. The more time that passes, the greater the possibility that a witness’s memory of a perpetrator will weaken.

**Ask Yourself:** How long was it between the crime and the lineup identification? Does this enhance or impair his/her accuracy?

In evaluating the reliability of a witness’s identification, you should also consider whether it was the result of a suggestive procedure. In that regard, you may consider everything that was done or said by law enforcement to the witness during the identification process including:

**(1) Lineup Composition:** A suspect should not stand out from other members of the lineup. The reason is simple: an array of similar-looking individuals forces witnesses to examine their memory. In addition, a biased lineup may inflate a witness’s confidence in the identification because the selection process seemed so easy to the witness. It is, of course, for you to determine whether the composition of the lineup had any effect on the reliability of the identification.
Ask Yourself: Did all or nearly all of the pictures/lineup members match the eyewitness’ description of the perpetrator? Does this strength or weaken the evidence against the defendant?

(2) Fillers: Lineups should include a number of possible choices for the witness, commonly referred to as “fillers.” The greater the number of possible choices, the more likely the procedure will serve as a reliable test of the witness’s memory. A minimum of six persons or photos, each of whom could reasonably be the perpetrator, should be included in the lineup.

Ask Yourself: How many reasonable candidates were in the lineup? Does this strength or weaken the evidence against the defendant?

In determining the reliability of the identification, you should also consider whether the identification procedure was properly conducted.

(a) Double-blind: A lineup administrator who knows which person or photo in the lineup is the suspect may intentionally or unintentionally convey that knowledge to the witness. That increases the chance that the witness will identify the suspect, even if the suspect is innocent. For that reason, whenever feasible, live lineups and photo arrays should be conducted by an officer who does not know the identity of the suspect.

Ask Yourself: Did the officer conducting the lineup do anything to suggest to the eyewitness which person or picture to choose, or not to choose? Does this strength or weaken the evidence against the defendant?
(b) **Instructions:** You should consider what was or what was not said to the witness prior to viewing a lineup or showup. Identification procedures should begin with instructions to the witness that the perpetrator may or may not be present and that the witness should not feel compelled to make an identification. The failure to give this instruction can increase the risk of misidentification. If you find that the police did or did not give this instruction to the witness, you may take this factor into account when evaluating the identification evidence.

**Ask Yourself:** Did the officer conducting the lineup tell the eyewitness that the perpetrator may or may not be present in the lineup? Does this strength or weaken the evidence against the defendant?

(c) **Confidence and Accuracy:** The witness made a confidence statement at the time the identification was made. Although some research has found that highly confident witnesses are more likely to make accurate identifications, eyewitness confidence is not a perfect indicator of accuracy.

**Ask yourself:** How much weight should I give the confidence statement?

(d) **Feedback:** Feedback occurs when police officers, or witnesses to an event who are not law enforcement officials, signal to eyewitnesses that they correctly identified the suspect. That confirmation may reduce doubt or produce a false sense of confidence in a witness. Feedback may also falsely enhance a witness’s recollection of the quality of his or her view of an event. It is for you to determine whether or not a witness’s recollection in this case was affected by feedback or
whether the recollection instead reflects the witness’s accurate perception of the event.

**Ask Yourself:** Did the police do anything that could inflate the eyewitness’ confidence?

**Does this strength or weaken the usefulness of the eyewitness’ confidence statement?**

Keep in mind that the presence of any single factor or combination of factors is not proof that a particular witness is incorrect. Instead, you should consider the factors that I have discussed as you assess all of the circumstances of the case, including all of the testimony and documentary evidence, in determining whether a particular identification made by a witness is accurate and thus worthy of your consideration as you decide whether the State has met its burden to prove identification beyond a reasonable doubt.
References


doi: 10.1002/acp.1796


