EYEWITNESS IDENTIFICATION JURY INSTRUCTIONS: DO THEY ENHANCE EVIDENCE EVALUATION?

by

MARLEE KIND BERMAN

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the Doctor of Philosophy, The Graduate Center, CUNY

2015
The manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the Dissertation requirements for the degree of Doctor of Philosophy

4/1/15
Date

(Print) STEVEN PENNAD
(Signature) 
Chair of Examining Committee

(Print) 
(Signature) 
Executive Officer

(Print) Dr. Jennifer Dysart
(Print) Dr. Deryn Strange
(Print) Dr. Edie Greene
(Print) Dr. Gary Wells

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

Mistaken eyewitness identifications are a leading cause of wrongful convictions. Even with procedural safeguards (e.g., cross-examination) in place, jurors still have difficulty evaluating the reliability of eyewitness identifications. The purpose of the present line of research is to examine whether the issue-specific judicial instructions, set forth by the New Jersey Supreme Court (*New Jersey v. Henderson*, 2011), effectively sensitize jurors to eyewitness identification accuracy.

Results of the first study indicate that the current *Henderson* instructions delivered on issues specific to a case are not as effective as intended. Results of the second study indicate a sensitivity effect, such that mock jurors were most sensitive to the quality of the eyewitness identification when they received judicial instruction on all eyewitness variables and evaluated the evidence prior to rendering a verdict. Implications of the research as well as ideas for future research are discussed.
Acknowledgments

I would like to thank my wonderful husband for his endless support throughout this process. Thomas, you have supported me and encouraged me for what seemed to be an endless number of years of schooling – I love you for this and for so much more.

Mom – had you not been on my case every day of my life, I’m not sure I could have gotten through 10 years of school. So thank you for the love and the nagging. Dad – yes, now you have to call me doctor.

And of course, this dissertation would not have been possible without the help from my advisor, Dr. Steven Penrod. Thank you, Steve!
Table of Contents

1. Introduction 1
   a. System variables 5
   b. Estimator variables 21
   c. Eyewitness ID reform 28
   d. Overview of current studies 34
   e. Study 1 35
      i. Method 36
      ii. Results 43
      iii. Discussion 48
   f. Study 2 54
      i. Method 55
      ii. Results 61
      iii. Discussion 69
   g. Overall Discussion
   h. References 80

2. Study 1 Tables 92
   a. Table 1 (Descriptive Statistics for Verdict)
   b. Table 2 (Effects of Step 1, Step 2, and Step 3 Variables on Verdict)
   c. Table 3 (Effects of Hypothesized Mediators on IVs and Interactions)
   d. Table 4 (Impact of Manipulated Variables on ID Accuracy)

3. Study 2 Tables 103
   a. Table 5 (Logistic Regression Overall Model)
   b. Table 6 (Verdict as a Function of Evid. Eval., Instr., & Evid. Qual.)
   c. Table 7 (Planned Comparisons of Interest Descriptives)

4. Study 2 Figures 105
   a. Figure 1 (Instruction Manipulation)

5. Study 1 Appendices 107
   a. Appendix A (No Eyewitness Instruction)
   b. Appendix B (Henderson Instruction)
   c. Appendix C (Post-Trial Questionnaire)

6. Study 2 Appendices 124
   a. Appendix D (No Eyewitness Instruction)
   b. Appendix E (Case-Specific Henderson Instruction)
   c. Appendix F (All-Inclusive Henderson Instruction)
   d. Appendix G (Post-trial Questionnaire) 137
Eyewitness Identification Errors and Suggested Safeguard Improvements

Although jurors tend to rely a great deal on eyewitness evidence (Penrod & Cutler, 1999), eyewitness identifications are not always accurate. Research indicates that the prevalence of mistaken eyewitness identifications is the primary cause for wrongful convictions in the United States (Gross, Jacoby, Matheson, Montgomery, & Patil, 2005; Scheck & Neufeld, 2006; Scheck, Neufeld, & Dwyer, 2001). In fact, according to recent data from the Innocence Project, eyewitness identification was the sole piece of evidence in many wrongful conviction cases (http://www.innocenceproject.org; The National Registry of Exonerations, 2014)

The United States Supreme Court initially examined the issues related to eyewitness identification in 1967, when several pivotal cases focusing on police misconduct during eyewitness identification procedures were decided (United States v. Wade, 1967; Gilbert v. California, 1967; Stovall v. Denno, 1967). In 1972 the U.S. Supreme Court determined that a judge should consider five factors when evaluating an eyewitness identification: 1) the opportunity for witnesses to view the criminal at the time of the crime, 2) the length of time between the crime and the later identification, 3) the level of certainty shown by the witnesses at the identification, 4) the witnesses’ degree of attention during the crime, and 5) the accuracy of the witnesses’ prior description of the criminal (Neil v. Biggers 1972). Also in Neil v. Biggers, the Court ruled that the per se exclusion rule, or the presumption that a judge could exclude unnecessarily suggestive identification procedures at trial, should be thrown out. This approach was affirmed in Manson v. Braithwaite (1977), wherein the Court argued that exclusion of eyewitness identification testimony should be based upon a two-step approach. In the first step, it must be determined whether there was in fact unnecessary suggestiveness in the identification procedure. If no evidence of suggestiveness is presented, the testimony should be admitted. If
there is evidence of suggestiveness, it must then be determined whether, regardless of the suggestiveness, the identification was in fact reliable. If it is determined, based on the *Biggers* criteria, that the identification is reliable (although also suggestive), the testimony should not be excluded (*Manson v. Braithwaite*).

The Manson two-step approach did not prove to be as successful as intended. In an article written 32 years after Manson was decided, Wells and Quinlivan (2009) argued that the Manson criteria have several weaknesses. The first problem lies within the first step – wherein one must show evidence of suggestibility. While some forms of suggestibility, such as inappropriate fillers or multiple viewings, are easy to determine, it can be nearly impossible to show evidence of whether the lineup administrator was biased or gave the witness hints (Wells & Quinlivan). Moreover, the criteria for determining reliability in the second step are not necessarily related to identification accuracy and the factors are largely based on self-report from the witness (i.e., certainty, view, and attention; Wells & Quinlivan). The issue with these self-reported variables is that research shows these variables can themselves be influenced by suggestive procedures, and are thus not independent of the suggestiveness evaluation (Wells & Bradfield, 1998).

More than thirty years after *Manson*, the New Jersey Supreme Court rendered a landmark decision in *New Jersey v. Henderson* (2011), one of many cases in which the reliability of an eyewitness identification was questioned. In *New Jersey v. Henderson*, the Court appointed a special master to hear from eyewitness identification experts and review the extant empirical research on eyewitness identification. Specifically, the Court reviewed research on both system variables (variables that are within the control of the state) and estimator variables (variables related to the perpetrator, environment, or witness; Wells, 1979). Due in part to the Special Master’s report, the New Jersey Supreme Court decided that the two-step approach for
determining eyewitness identification admissibility was not adequate and based their decision on three factors: (1) it contained no sufficient reliability measure, (2) it was an inadequate deterrent for inappropriate police conduct, and (3) there was over-reliance on the jury’s capacity to evaluate the reliability of the identification. Instead, the Court determined that, if the defendant can show evidence of suggestibility, all system and estimator variables relevant to the identification should be investigated at a pretrial hearing (New Jersey v. Henderson, 2011). If the prosecution demonstrates that the identification is still reliable, the identification should be admitted. Upon admittance of the identification, however, courts must instruct the jury on the research concerning relevant system and estimator variables in order to help jurors evaluate the reliability of the identification (New Jersey v. Henderson, 2011). The Henderson court proposed that “even with matters that may be considered intuitive, courts provide focused jury instructions” because “it is the court’s obligation to help jurors evaluate evidence critically and objectively to ensure a fair trial” (p. 123-124). Specifically, the Henderson court noted the added benefits of these enhanced jury instructions, such that 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” (p. 126).

Below are sample instructions for two system variables (lineup instructions; feedback) and two estimator variables (weapon focus; duration) drafted by the Criminal Practice Committee and the Model Criminal Jury Charge Committee in response to Henderson:

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:

**Instructions:** You should consider what was or what was not said to the witness prior to viewing a photo array. Identification procedures should begin with instructions to the witness that the perpetrator may or may not be in the array 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/did not] give this instruction to the witness, you may take this factor into account when evaluating the identification evidence.

**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.

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:

**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.

**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.

Although New Jersey has adopted these new standards, researchers have yet to thoroughly investigate whether the *Henderson* instructions are actually an effective method for sensitizing jurors to eyewitness identification reliability. Sensitivity in this context refers to an improvement in jurors’ abilities to differentiate between conditions in which eyewitness identification evidence is more versus less likely to be accurate and reaching verdicts accordingly (Cutler, Dexter, & Penrod, 1989). Alternatively, the presence of judicial instructions may reveal evidence of a skepticism effect (Cutler, Dexter, & Penrod, 1989), wherein jurors, instead of discerning between reliable and unreliable identification, become skeptical of
identifications in general. Research on the efficacy of earlier versions of eyewitness judicial instructions suggests instructions may not be an effective safeguard (see Berman & Penrod, unpublished). It should be noted that in a more recent decision, the Oregon Supreme Court ruled that it is the prosecution’s burden to show that the identification is adequately reliable (Oregon v. Lawson, 2012).

Below is a review of the literature on a number of factors known to affect the accuracy of eyewitness identifications as well as the judicial instruction safeguard intended to educate jurors about factors that may affect the accuracy of an identification, should unreliable eyewitness identification be admitted into court. This review is not exhaustive but is intended to provide an overview of eyewitness factors employed in new research reported below. The new research is designed to test whether the Henderson instructions do in fact sensitize jurors to variations in trial evidence associated with factors known to influence eyewitness identification reliability.

**System Variables**

System variables refer to variables that are ultimately in the hands of the criminal justice system (e.g., showup vs. lineup, multiple viewings, biased lineup instructions, confirmatory feedback; Wells, 1978). These are variables that, depending upon how they are administered, may affect the accuracy of an eyewitness identification. Moreover, system variables are easily controlled in both experimental and real-world settings and may influence witness decisions.

Dating back to the mid 1900s, social psychologists have looked at the ways in which people are unduly influenced by the people around them, particularly normative and informational social influences (Deutsch & Gerard, 1955). Normative and informational social influences refer to the idea that people may conform to what they see as a social norm or believe
what they believe others expect them to believe. These normative and informative social influences, among other social influences, are often shown to play a powerful role in the eyewitness identification procedure. Thus, decades of research have looked at whether eyewitnesses are influenced by the lineup administrator.

Lineup administrators can influence eyewitnesses, through both verbal and nonverbal behaviors, in a variety of ways. For example, administrator knowledge of the suspect may influence the identification (e.g., Buckhout, 1975; Wells, 1988). In fact, lineup administrators can be quite influential if they have knowledge of the suspect and are determined to obtain a positive identification. In one study, prior to the lineup, all participants heard the same recording of standard, unbiased lineup instructions (Clark, Marshall, & Rosenthal, 2009). Researchers manipulated whether, after twelve-seconds, the lineup administrators kept silent (no influence), made a cautionary statement (e.g., “take your time”, “there’s no rush”) to the witness (subtle-influence), or reminded the eyewitness to choose the person who was most similar to the perpetrator (similarity-influence; Clark, et al.). For analysis, a distinction was made between the target-absent lineups, in which there was a designated innocent suspect (DES: determined prior to the study) and a worst-case scenario innocent suspect (WCS: determined based on which innocent suspect is identified most often). Results mainly compared the effects of both subtle and similarity-influence conditions to the no influence conditions, respectively. In the former comparison, the overall identification rate as well as the false identification rate (for the designated innocent suspect) was greater in the subtle-influence than the no-influence condition (Clark et al.). The major differences between the two types of target-absent lineups were that foil identification rates were lower while false identification rates were higher in the WCS assumption. In the latter comparison, the overall identification rate was higher in the similarity-
influence conditions for both target-present and target-absent lineups (Clark et al.). Additionally, there was a greater rate of correct identifications in the target-present conditions in the similarity influence versus no influence (Clark et al.). Overall, the various types of lineup administrator influences did yield differences in the pattern of eyewitness identification decisions, none of which increased the probative value of the identifications (Clark et al.).

In a more pronounced demonstration of the effects that a non-blind administrator can have, Alberts, Duncan, Wallace and Penrod (2008) trained administrators to ”steer” witnesses to make positive identifications of suspects (both guilty and innocent suspects) and avoid identifications of non-suspects (foils) and to accomplish this steering without arousing the suspicion of witnesses. Steering was highly successful: selection rates for “steered” witnesses were: 57% guilty suspects, 32% innocent suspects and 19% filler identifications versus 18% guilty suspects, 10% innocent suspects and 39% filler identifications. Witnesses were essentially unaware that they had been influenced. Clark, Brower, Rosenthal, Hicks, and Moreland (2013) used similar procedures and obtained similar results. Clark et al., (2013) trained participants to administer a lineup, half of whom were trained to obtain positive identifications of a suspect. Identification rates of both guilty and not guilty suspects were higher when the lineup administrators led witnesses to make an identification (Clark et al., 2013), indicating the extent to which lineup administrators exert influence over witnesses.

In an examination into how and why this influence occurs, Greathouse and Kovera (2009) found an increase in suspect identifications when the administrator was single blind, provided biased instructions, and administered a simultaneous lineup. Moreover, they determined that double-blind lineups were more than two times as diagnostic as the single-blind lineups. Diagnositicity, which was calculated as the proportion of identifications of the suspect in target-
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

present lineups divided by the proportion of identifications of the suspect in target-absent lineups (Wells & Lindsay, 1980; Greathouse & Kovera), refers to the extent to which one can use the suspect identification as a reliable indicator of identification accuracy.

Showup versus Lineup

Showups, which are a common form of eyewitness identification, can occur in a nearly unlimited number of places and courts have worried that they can be quite suggestive. In Stovall v. Denno, in which the identification was made via a showup, the Supreme Court opined that it is “hard to imagine a situation more clearly conveying the suggestion to the witness that the one presented is believed guilty by the police” (Stovall v. Denno, 1967). After a series of studies (two experiments and one archival study), Gonzalez, Ellsworth, and Pembroke (1993) proffered that showups are, in fact, no more suggestive than lineups. Specifically, the authors found no difference in rates of misidentifications between lineups and showups in their lab experiment (Gonzalez et al., 1993). In another study, the rate of false identifications did not differ significantly between showups and lineups (Yarmey, Yarmey, & Yarmey, 1994). After correcting for guessing, however, accurate identification rates were greater when participants were shown a six-person compared to a one-person photographic lineup, (Yarmey et al.).

One of the advantages of utilizing a showup over a traditional multi-person lineup is that it can be conducted quickly, since memory has been shown to decrease over time (Light, 1996). Clark and Godfrey (2009), found a decrease in identifications of guilty suspects over time. Furthermore, in their systematic review of the identification literature, Clark and Godfrey found no significant difference between lineups and showups in the comparison between correct identification rates from target-present conditions foil identifications from target-absent conditions.
In a meta-analytic comparison of eyewitness accuracy rates in showup and lineup presentations, Steblay, Dysart, Fulero, and Lindsey (2003) examined twelve tests of the hypothesis across eight studies. Results indicated that eyewitnesses exhibited generally higher rates of correct decisions (identifications of targets when present and rejections of arrays when the target is absent) when given a showup compared to a lineup ($r = -.18$; Steblay et al., 2003). When they did choose a suspect, eyewitnesses in showup conditions correctly rejected the array more often than eyewitnesses in lineup conditions ($r = -.32$), but the incorrect identifications in lineups were distributed across multiple innocent foils with the result that there was no difference in false identification rates between showups and lineups (Steblay et al.). However, as Steblay, et al. note, if the arrays used in these studies had been fair arrays, the false identification rate would have been 7.2% versus the 15% rate for showups. In a more recent study of undergraduate students in a mock crime scenario, Dickinson (2006) concluded that showups were not more suggestive than lineups, although there are two exceptions. Results showed that a showup was more suggestive than a lineup when the showup was conducted immediately following the crime and the innocent suspect’s clothing was similar to that of the perpetrator (see Dysart, Lindsey, & Dupuis, 2006) and when an innocent suspect appears in both a showup and a lineup (Dickinson). The latter is unsurprising, as having a suspect appear in multiple identification procedures has been shown to increase the suggestiveness of the procedure(s). The finding that short retention intervals and clothing similarity foster suggestiveness underscores one of the major weaknesses of most experimental comparisons of one-person versus multi-person presentations, namely that the one-person conditions lack most of the suggestive real-world show-up characteristics which have concerned courts.
A more recent experiment (Wetmore, et al. 2015) compared showups and lineups on immediate tests and tests conducted after 48 hours. The authors report: “Participants (N = 1486) viewed a mock-crime video and then were presented with a showup or a simultaneous lineup, either immediately or a 48 h delay. Receiver operating characteristic (ROC) analyses revealed that a showup never resulted in better identification accuracy than a lineup. Their results are illustrated in the following figures…in which the desirable outcome is the represented by lines which are closest to the upper left corner of the figure (higher correct ID rates coupled with low mis-ID rates).

Figure: ROC curves comparing showups (SU) and lineups (LU; top panel), showups and lineups at the different retention intervals (middle panel), and the lineups separated by fairness (bottom panel). The thick diagonal line in each panel indicates chance performance.

Wetmore, et al. conclude: “lineups were more diagnostic than showups. This was true when showups were compared to lineups immediately and after a 48 h delay. These results replicate previous work comparing these identification procedures (Gronlund et al., 2012)… It also was true that both fair and biased lineups resulted in better diagnostic accuracy than showups…”

**Biased Lineup Instructions**

Several factors have been shown to increase the suggestiveness of a lineup, one of which is the initial instructions (or lack thereof) given to the witness by the lineup administrator. A common operationalization of biased lineup instructions in the literature is the mention, or lack
of mention, that the suspect may or may not be in the lineup. In one of the earliest studies looking at the effect of biased lineup instructions, Malpass and Devine (1981) manipulated the lineup instructions given to the witness. In the biased instruction condition, the police informed the witness that they believed the perpetrator was present in the lineup. In the unbiased instruction, however, the police informed the witness that the perpetrator may be in the lineup, but it was possible that he was not in the lineup. Witnesses in the unbiased instruction condition were given the option to choose no suspect, whereas the witnesses in the biased instruction were forced to choose one of the people in the lineup. Unbiased lineup instructions dramatically lowered the rate of target-absent identifications (from 78% to 33%) but did not reduce the rate of correct identifications (biased = 75%, unbiased = 83%; Malpass & Devine).

In one meta-analysis, 18 studies were included in an investigation into whether biased lineup instructions increase witness choosing rates and decrease the accuracy of identifications, compared to unbiased lineup instructions (Steblay, 1997). Overall, Steblay found that biased instruction does significantly affect eyewitness identification accuracy. Specifically, witnesses who were given biased instructions chose someone from the lineup significantly more often than witnesses who were given unbiased instructions. The type of lineup (i.e., target-absent versus target-present), however, modified the accuracy of eyewitness identifications, such that performance was affected most when eyewitnesses were given biased lineup instructions for a target-absent lineup (unbiased = 60%, biased = 35%; $r = .28$; Steblay). There was no difference in eyewitness identification accuracy, however, when participants saw a target-present lineup (unbiased = 54%, biased = 53%). Additionally, participants were more confident in their identifications when they received biased instructions and were given a target-present lineup (Steblay).
In part, based on the assumption that eyewitness confidence may influence the eyewitness’ presumed identification accuracy in target-present lineups, Leippe, Eisenstadt, and Rauch (2009) showed participants a 5-minute video of a theft, subsequently administered a set of questions regarding participants’ memory of the video, and administered either a target-absent or target-present lineup. Leippe et al., manipulated witness confidence prior to the identification by providing either positive, negative, or no feedback regarding the participants’ memory (compared to other witnesses’ memories). The researchers also manipulated whether participants received biased or unbiased instructions as well as whether the lineup was target-present or target-absent (Experiment 1). In Experiment 2 the methodology was nearly the same as Experiment 1; however, there was no memory feedback task and the type of lineup was manipulated such that participants viewed a target-present with either dissimilar or similar foils or a target-absent with either dissimilar or similar foils. Across experiments, participants who received biased instructions had a higher rate of choosing and a higher rate of false identification of the similar foil compared to those who received unbiased instructions (Leippe et al.). Results of the first experiment showed that witnesses were more confident when they received biased instructions and positive memory feedback (Leippe et al.). Results of experiment 2 indicated that the presence of biased instructions increased witness confidence when the thief was present (regardless of foil type) as well as when the thief was absent and the foils were dissimilar (Leippe et al.). Based on these two experiments, it seems that both witness identification confidence and witness identification accuracy increase when a target stands out in a lineup and the witness receives biased lineup instructions.

To further investigate the idea that providing witnesses with pre-identification feedback may have an effect on witnesses’ identification decisions, Quinlivan et al. (2012) manipulated
pre-admonition suggestions. That is, along with manipulating whether participants received biased or unbiased instructions, Quinlivan et al. manipulated whether, prior to the instructions, the administrator suggested that the perpetrator was likely in the lineup (i.e., “I could really tell you were paying a lot of attention; surely you are going to be able to pick the person out of the lineup”). When the instructions were unbiased, there was an increase in false identifications and overall choosing rates when participants received the pre-admonition suggestions compared to no pre-admonition suggestions (Quinlivan et al.). No significant differences in false identifications or choosing rates were found in the biased instruction conditions as a function of pre-admonition suggestions (Quinlivan et al.). Results of this study demonstrate that unbiased instructions may be more effective than biased instructions, but only when the administrator actually refrains from suggesting that the perpetrator is likely in the lineup. That is, unbiased lineup instructions can only be effective when the lineup administrator does not make statements to contradict those instructions.

**Confirmatory Feedback**

It can be difficult to control what lineup administrators say to witnesses in a real-world setting - the biased-instruction research demonstrates the effect of biased pre-identification communications on witness accuracy, but the suggestiveness of a lineup (or showup) can also increase after the identification is made if the administrator informs the witness that they have correctly identified the suspect. A witness who is unsure at the time of the identification may appear to be quite certain of their identification when it comes to subsequent hearings and the trial. This increase in eyewitnesses’ certainty can sometimes be linked back to the feedback that the person conducting the lineup gives to the witness, or confirmatory feedback (Wells, 1998). While a biased lineup instruction has been shown to influence whether or who the witness
chooses, confirmatory feedback can influence how confident the witness is after his/her 
identification. If the identification is inaccurate and the witness is confident that they made a 
correct identification, there is obviously a problem.

In an early study on the feedback effect, participants watched a video of a man robbing a 
store with a handgun (Wells & Bradfield, 1998). Participants were shown a photo lineup and 
asked to identify the suspect from the video. Researchers manipulated whether participants 
received confirming feedback (informing participants that they did identify the suspect), 
disconfirming feedback (informing participants that they did not identify the suspect and then pointing out the actual suspect), or no feedback (Wells & Bradfield). Following the identification, 
participants who received feedback were significantly more certain in their identification, their 
view of the perpetrator, their ability to perceive facial details, their attention to the event, 
williness to testify, trust in their identification under the conditions, and more (see Wells & 
Bradfield for complete list of effects). Significant effects appeared in the comparison between 
the confirming and disconfirming feedback conditions across (average \( d = 1.08 \) across 10 
measures); as well as in the comparison between the confirming and no feedback conditions 
(average \( d = .75 \) across 10 measures; Wells & Bradfield).

In quite a short period of time, researchers conducted many studies looking at the effect 
of feedback on eyewitnesses. In 2006, Douglass and Steblay conducted a meta-analysis of ten 
published and four unpublished experimental studies in which confirmatory feedback was 
investigated. In all studies included in the analysis, participants viewed a staged crime video and 
were then instructed to make identifications. In addition to a comparison of feedback groups to 
no feedback groups, a comparison of confirming feedback and disconfirming feedback was 
included in a secondary analysis. Overall, the effect of post-identification feedback on witness
certainty was large, based on Cohen’s $d = 0.79$; Douglass & Steblay, 2006). Moreover, participants who received confirmatory feedback exaggerated reports of greater witnessing conditions as well as better memory (at time of lineup and in general) compared to those who received either no feedback or disconfirming feedback (Douglass & Steblay). In a more recent meta-analysis, Steblay, Wells, Douglass, and Bradfield (2014) found that the effect of post-identification feedback has remained constant over time (mean effect size for studies after 2007: $d = .73$, $k = 11$; mean effect size for studies before 2007: $d = .70$, $k = 8$).

Douglass and colleagues were interested in determining the effects of post-identification feedback on peoples’ evaluations of the witness testimony and the identification procedures (Douglass, Neuschatz, Imrich, & Wilkinson, 2010). To that end, they conducted two experiments in which evaluators rated witnesses. In the first experiment, evaluators viewed the testimony of an eyewitness who, after their identification, either received confirming feedback, disconfirming feedback, or no feedback. In the second experiment, evaluators viewed an identification procedure wherein witnesses received confirming feedback and were either told or not told to disregard that feedback (Douglass et al. 2010). Across experiments, evaluators rated eyewitnesses who received confirming feedback more positively than eyewitnesses who received either disconfirming or no feedback (Douglass et al., 2010). Additionally, instructions to disregard the feedback (in Experiment 2) did not affect witness evaluations (Douglass et al., 2010). Taken together, this research indicates that witnesses who receive confirming feedback, which in turn increases their confidence, are perceived as more accurate, even when they are not. In another study, Douglass and Jones (2013) found that jurors who were shown a video of the identification procedure were more sensitive to increases in eyewitness confidence from the time of the identification to the time of the trial.
Gurney and colleagues (2013) took the malleability research one step farther and tested whether positive or negative non-verbal feedback would affect witness confidence. Participants watched CCTV footage of a crime scene and answered a series of questions. While answering questions the participants given either affirming feedback (a head nod), disconfirming feedback (a head shake) or no feedback. Those presented with affirming non-verbal feedback produced inflated confidence judgments compared with those presented disconfirming non-verbal feedback and this was true regardless of their accuracy.

Smalarz and Wells (2014a) investigated other ways in which confirming postidentification feedback can affect an identification. After participants received either confirming feedback or no feedback, the experimenter pretended to have given the witness the “wrong” lineup, and subsequently gave the witness the “correct” lineup. Results indicated that feedback impaired participants’ ability to identify the culprit in a final memory test (Smalarz & Wells, 2014a). In another study, Smalarz and Wells (2014b) found that participant-evaluators were able to distinguish between accuracy and credibility of participant-witness’ videotaped testimony when participants-witnesses were not given feedback. When participant-witnesses were given feedback, however, there was no difference in evaluators’ ratings as a function of accuracy/credibility (Smalarz & Wells, 2014b).

One proposed improvement to this feedback effect is to employ blind lineup administrators. To test the hypothesis that feedback from lineup administrators who know the identity of the suspect may bias eyewitnesses, Dysart, Lawson, and Rainey (2012) tested whether feedback from lineup administrators, who are blind to the identity of the suspect, have the same influence on eyewitness confidence as non-blind administrators. Participants in this study watched a video of a mock crime and were subsequently asked to identify the suspect from either
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

a target-present or target-absent lineup (Dysart et al., 2012). Researchers manipulated whether the participants did or did not receive feedback from an administrator that was presumed blind or presumed non-blind. Post-identification feedback did have an effect on participant ratings of confidence and other judgments but only when participants received feedback from an administrator that was in the room during the video (presumed non-blind). There was no effect of feedback on participant ratings of confidence and other judgments, however, when the lineup administrator was not in the room during the video (presumed blind). Thus, as the feedback effect diminishes, eyewitness reliability presumably increases. This finding presents a compelling potential safeguard to this feedback effect - double-blind lineup administration.

Multiple Viewings

According to an archival analysis of criminal investigations, repeated identification procedures are relatively common (Behrman & Davey, 2001). For instance, a witness may be shown a lineup containing a suspect’s photo and subsequently also be shown a lineup containing the suspect’s photo with different fillers, at which point the witness is likely going to notice the commonality (i.e., the suspect) between the two lineups. If the witness then identifies the suspect, how can we be sure that he/she did not simply recognize the suspect from the previous lineup as opposed to from the crime scene? Such repeated eyewitness identification procedures are considered so suggestive, in fact, that defense attorneys are quite familiar with the photobiased lineup (Deffenbacher, Bornstein, & Penrod, 2006) and most experts on eyewitness testimony consider it to be generally unreliable (Kassin, Tub, Hosch, & Memon, 2001). Early research implementing a photobiased lineup showed that participants tended to remember faces they had seen before, more so than where/when they came across those faces (Brown, Deffenbacher, Sturgill, 1977). In a meta-analysis, Deffenbacher et al., (2006) investigated the hypothesis that
witnesses’ prior exposure to mugshots increases the rate of false identifications and decreases the rate of correct identifications. Overall, eyewitnesses are more likely to choose a suspect whose mugshot they have already seen, compared to a novel suspect ($Z_{ma} = 2.33$; Deffenbacher et al.). Moreover, this effect was greatest when studies in which the actual suspect was included in the lineup were excluded from the analyses ($Z_{ma} = 4.91$; Deffenbacher et al.). In fact, in an investigation into the suggestibility of showups, Dickinson (2006) found that when an innocent suspect appears in both a showup and a lineup, the chances of that suspect being identified increases substantially.

Merely informing an eyewitness that he/she will get the chance for a second viewing may be enough to affect the accuracy of the identification (Duckworth & Kreiner, 2009). In a computer task, researchers manipulated whether or not participants were told that they would be given a second chance to view a sequential lineup via a randomly-ordered lineup administered on a computer (plus a control condition in which participants viewed the lineup only once). All participants were briefed on the difference between simultaneous and sequential lineups and participants in the “Told” condition were subsequently informed that they would have a chance to view the lineup twice (all instructions were given prior to the first lineup; Duckworth & Kreiner). Overall, participants who viewed the lineup twice were more likely to make an incorrect identification than those who viewed the lineup only once (Duckworth & Kreiner). Compared to those who were told about the second viewing, participants who were not told that they would have a second viewing made correct identifications of the perpetrator more often and made incorrect rejections of the lineup less often (Duckworth & Kreiner).

In another study, researchers conducted two experiments to measure the differential effects of multiple identification procedures when the time between the first and second
identification is manipulated (Experiment 1: one week delay; Experiment 2: 30 minute delay; Godfrey & Clark, 2010). Following a robbery video, participants were asked to write down anything they remembered about the crime and the perpetrator along with their projected confidence in their ability to identify the perpetrator. Researchers then administered either a photo showup or gave participants a filler task. In the first experiment, participants returned to the laboratory the next week at which point they were presented with either a guilty-suspect or innocent-suspect lineup (Godfrey & Clark). In the second experiment, participants were given another filler task to complete and were thus presented with the lineup only thirty minutes after the showup. When the interval was short (30 minutes), few participants changed from a nonidentification to a suspect-identification between showup and lineup, compared to nearly half of the participants who made this change when the interval was long (one week; Godfrey & Clark). When the interval was short, there was an increase in foil identifications from showup to lineup, but there was no change in correct or false identifications. Conversely, when the interval was long, there was an increase in both correct and false identifications from showup to lineup, but no change in foil identifications. Thus, although suspects (both guilty and innocent) are more likely to be identified in multiple identification procedures, the identification holds little probative value (Godfrey & Clark). Additionally, across experiments, consistent witnesses were more confident than inconsistent witnesses and confidence rates for all participants increased between showup and lineup when the interval was short (Godfrey & Clark). Thus, witness confidence is not necessarily a sound indicator of eyewitness accuracy.

Recent research further indicates the problematic effects of showing witnesses lineups and/or showups on multiple occasions. In one study, participants watched a video of a crime and were asked to identify the culprit (either guilty or innocent) from two lineups of the same format
(either sequential or simultaneous), with a two-week time period between lineups (Steblay, Tix, & Benson, 2013). Results supported the researchers’ hypothesis that participants who made choosing errors at the first lineup would continue, rather than correct, their identification error at the second lineup. Moreover, there was no difference in confidence among those who made correct versus false identifications, and choosers were significantly more confident than non-choosers (Steblay et al., 2013).

Another important issue regarding multiple viewings is whether or not witnesses are granted the opportunity to see each of the photos in a sequential lineup more than once. Horry, Brewer, Weber and Palmer (2015) tested the effects of two laps (either required or optional) of a sequential lineup on witness’ identification decisions. Choosing rates (of both fillers and culprits) increased in the second lap (40% changed original response, mostly from nonidentification to identification), but identification accuracy did not differ as a function of number of laps (Horry et al., 2015). In another study, Abraham (2012) found that overall accuracy was greater when participants viewed each mugshot once (47.6%), compared to two laps (19.7%) and two laps with a warning about the potential negative effects of multiple viewings (31.1%).

Backloading, or allowing the witness to believe the sequential lineup contains more photos than it truly does, may be an effective way to decrease rates of false identifications and encourage more conservative responses. Horry, Palmer, and Brewer (2012) showed participants a sequential lineup (either target-present or target-absent) that was either non-backloaded (6 photos) or backloaded (either 12 or 30 photos). The researchers also manipulated whether the suspect appeared in the beginning (position 2) or later (position 6) in the lineup. When the lineups were backloaded, the innocent suspects in position 6 were chosen less frequently compared to when the lineups were non-backloaded (Horry et al., 2012). Moreover, backloading
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

the lineup deterred witnesses from becoming more lenient as the lineup progressed, which was indicated by an increase in witnesses’ choosing rates from position 3 to position 5 when the lineup was not backloaded (Horry et al., 2012).

Taken together, research on system variables is useful and the variations in the variables can be easily controlled for in an experimental setting. Indeed, it has been argued that researchers should put their efforts into investigating system variables over estimator variables, because estimator variables are often subjective in nature and measured after the fact (Wells, 1978).

Estimator Variables

Estimator variables refer to those variables that are completely out of the hands of the criminal justice system, such as, viewing time (the amount of time the witness had to view the perpetrator), weapon focus (whether or not there was a weapon to draw the eyewitness’ attention), disguise (whether or not the witness was wearing a disguise), time lapse (the amount of time between the event and the identification; Wells, 1978). Similar to system variables, estimator variables are easily controlled for (or manipulated) in experimental settings. Thus, while control over these variables in real-world settings is not feasible, it is possible to estimate eyewitness accuracy based on reports.

Viewing Time

The amount of time in which a witness is exposed to the perpetrator is one factor that is likely to have an effect on the witness’ identification accuracy. In a meta-analytic review of over 150 facial identification studies and over 35 eyewitness identification studies, Shapiro and Penrod (1986) looked at the inclusion of exposure time as a study characteristic and/or an independent variable. Specifically, exposure time was defined as the amount of time (measured in seconds) in which subjects were able to view each target. Total exposure time was defined as
the total amount of time (measured in minutes) in which subjects were able to view all targets. Results of this meta-analysis indicated that as exposure time increased, participants were more sensitive in their eyewitness identification, that is there were higher hit rates ($d = 0.61$) and lower false alarms ($d = 0.22$) when exposure time was longer compared to when exposure time was shorter (Shapiro & Penrod).

An analysis of 640 eyewitness identification attempts across 314 actual police lineups in London revealed that witnesses who were exposed to the perpetrator for more than one minute were significantly more likely to accurately identify the perpetrator, compared to witnesses who were exposed to the perpetrator for less than one minute (Valentine, Pickering, & Darling, 2003). In another study, exposure time was manipulated in a simulated crime such that participants saw the perpetrator for either 12 seconds or 45 seconds (Memon, Hope, & Bull, 2003). Participants were subsequently shown either a target-present or a target-absent lineup and asked to identify the perpetrator. Irrespective of the age of the participant, those in the long exposure (45 seconds) condition were significantly more accurate than those in the short exposure (12 seconds) condition (Memon et al.).

Whereas researchers have mostly tested larger ranges of exposure times, based upon previous conclusions (e.g., Montejano 2005; Moore, 2006, Tietjen, 2005), an eyewitness to a shooting incident may have less than one second to view the perpetrator (Sharps, Janigian, Hess, & Hayward, 2009). Thus, it is important that research include various exposure times to account for the possible ranges in exposure time in real world situations. In one study, undergraduate students viewed a simulated, digital crime scene and were subsequently asked to describe what they had seen (Sharps et al.). Among other manipulated variables (i.e., gender of perpetrator and presence of a weapon) the researchers examined the effect of exposure time. Participants were
randomly assigned to view the crime scene for either .5 seconds, 2 seconds, or 5 seconds (Sharps et al.). Results showed that the length of time participants were exposed to the crime scene significantly affected several dependent variables. Specifically, participants incorrectly judged the gender of the perpetrator, the type of weapon, and the presence of a weapon in the shorter exposure condition (.5 seconds) compared to the medium (2 seconds) and long exposure (5 seconds) conditions (Sharps et al.). Surprisingly, participants in the long exposure condition, compared to both the medium and short exposure conditions, misidentified the race of the perpetrator significantly more often (Sharps et al.). These results indicated that witnesses may make considerable errors in identification of a perpetrator, depending upon the amount of exposure time; however, the range of exposure times used in this particular study may not be representative of many eyewitness identification scenarios.

In a more recent meta-analysis, the effect of exposure time on the accuracy of eyewitness identification was analyzed using a sample of thirty-three effect sizes from twenty-five articles (Bornstein, Deffenbacher, Penrod, & McGorty, 2012). The reported range of the differences between exposure time (between long and short duration) was .7 seconds to 3570 seconds. Results of this meta-analysis indicated that participants in the longer exposure conditions made significantly more accurate identifications compared to participants in the shorter exposure conditions, and the increased exposure time effect was non-linear ($r = .30$; Bornstein, et al.).

**Weapon Focus**

Another estimator variable that can affect the reliability of an identification is the presence of a weapon, which can divert the witness’ attention. Thus, if an eyewitness reports noticing a weapon present, they will theoretically be less able to accurately identify the perpetrator than if the weapon was not present. In an early meta-analysis of 12 studies (19 tests),
Steblay (1992) found that although more than two thirds of the tests did not show a significant difference between conditions in which the weapon was present versus absent, there was a reliable effect. In a study described earlier, in which exposure time had an effect on the accuracy of the eyewitness identification, there was no effect of the presence/absence of a weapon (Valentine, Pickering, & Darling 2003). The results of this study should be noted with caution, however, as the identifications were naturalistic and there was thus no way to establish the ground truth. In another study, researchers looked at eyewitness identification reliability as a function of sex differences and weapon focus (Shaw & Skolnick, 2001). Results showed that there was an own sex identification bias; however, there was no weapon focus effect and no interaction between sex differences and weapon focus (Shaw & Skolnick).

Some researchers have opined that the weapon focus effect is due to the presence of an unusual item or stimulus, not specifically a weapon. To test this hypothesis, in one study, college students viewed a video of a scene in which a perpetrator had a gun, celery, or neither (Mitchell, Livosky, & Mather, 1998). In the first experiment, the participants in the celery condition had a less accurate memory for the details of the perpetrator compared to participants in both the gun and control conditions. In the second experiment, results were consistent with a weapon focus effect, in that participants’ memory for the details of the perpetrator and the scene were less accurate when the perpetrator had a gun, compared to no item (Mitchell et al.). In another study, across two experiments, the threat level and unusualness of the object was manipulated, along with the presence or absence of the object in the context of a videotaped hair salon scene or electronic shop scene (Experiment 1 object was either: a pair of scissors, handgun, man’s wallet, or a raw whole chicken; Experiment 2 object was either: a screwdriver, a butcher knife, a pair of sunglasses, or a toy figurine; Pickel, 1998). Regardless of threat, participants had more difficulty
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

remembering details of the scene in which the perpetrator had an unusual object (Pickel). Thus, a more accurate description of the weapon focus effect may be that people will focus on unusual or unexpected objects (including, but not limited to, a weapon) and subsequently pay less attention to other details of the crime scene or perpetrator. In a further test of this unusual item hypothesis, Hope and Wright (2007) simultaneously showed participants slides of a man entering a store with either a gun, feather duster, or wallet in hand, as well as a rapid presentation of even numbers. The presentation was set up such that an odd number was presented at the same time as either the gun (weapon), feather duster (unusual object), or wallet (control). The researchers found that participants who saw the control object were able to detect more of the man’s features compared to both the unusual object and weapon conditions (Hope & Wright).

In one study, researchers found that the weapon focus effect disappeared after participants were instructed a priori that they should not focus on the weapon but instead on the perpetrators’ physical appearance (Pickel, Ross, & Truelove, 2006). While this finding is interesting, it is not very practical. Specifically, eyewitnesses are not likely to be informed that they should pay attention to the perpetrator in a real-life scenario.

Many studies have been conducted, however, in which there is empirical support for the weapon focus effect. In one study, across two experiments, participants were shown a series of slides in which a customer either hands a check to the cashier or pulls out a gun on the cashier (Loftus, Loftus, Messo, 1987). Researchers used an eye movement tracker to record the amount of time participants looked at either the check or the weapon. Results showed that participants looked at the weapon, compared to the check, for a longer period of time. Furthermore, participants in the weapon condition had less accurate memories for the details of the crime (Loftus et al.). Another consideration is that the effect of weapon focus may be more substantial...
when the weapon is completely, versus partially, visible. In one study in which participants viewed a mock crime scenario, those who saw a highly visible weapon had significantly poorer recall of details compared to those who saw only a partially visible weapon (Kramer, Buckhout, & Eugenio, 1990). A recent dissertation study by one of our students at John Jay College (an active duty Connecticut police chief; DeCarlo, 2010) illustrates the effect. DeCarlo used a videotaped robbery and found that in a no weapon condition, witnesses were able to correctly identify the target 78% of the time. When a weapon was implied by the perpetrator waving his hands around in his pocket, accuracy dropped to 55%. When a weapon was actually shown, accuracy dropped to 33%. When the perpetrator was absent from the lineup the correct rejection was 89% for the no weapon condition, but when a weapon was implied, accuracy dropped to 76% and when a weapon was actually shown, correct rejections dropped to 65% (DeCarlo, 2010).

Fawcett, Russell, and Peace (2013) conducted a meta-analysis to examine the ways in which the inconsistent presence of the weapon focus effect in the research may be a function of methodological differences between studies. Results demonstrated an overall effect of weapon focus (g = .53), characterized by a decrease in witness memory in the presence of a weapon or unusual object, that was moderated by several variables but unaffected by the type of study (either real-world, simulation, or laboratory; Fawcett et al.). Specifically, the weapon focus effect decreased as retention interval increased (immediate identification: g = .91; short duration: g = .66; long duration: g = .38). Additionally, the weapon focus effect increased when the threat level was high (g = .76) compared to low (g = .37; Fawcett et al.). Thus, while the actual type of study did not affect the presence of the weapon focus effect, it seems that methodological variations did cause significant fluctuations in the results, which (together with under-powered designs) more or less explains the inconsistency across studies.
**Time Delay (or Retention Interval)**

Research dating back to the late 1800’s and early 1900’s suggests that memory decays over time (Ebbinghaus, 1908). This memory decay is of particular importance to eyewitness identification researchers, because there is often a significant amount of time between the crime and when the witness makes an identification. This amount of time between when the witness views the perpetrator and when they identify the perpetrator is often referred to as the “retention interval.” It is expected that witnesses’ memories may decay over time, such that they are less accurate in their identification after more and more time passes. Thus, the retention interval is another very common variable manipulated in experiments on eyewitness identification.

In an early meta-analysis, Shapiro and Penrod (1986) analyzed 38 eyewitness identification studies in which retention interval was among the manipulated variables. In general, longer retention interval had a reliable impact on the accuracy of eyewitness identification \((r = -.29; \text{Shapiro & Penrod})\). The retention interval variable, however, which is continuous in nature, was coded as a dichotomous variable (i.e., short versus long retention interval) such that the total variance could not be completely accounted for (Shapiro & Penrod). In a more recent meta-analysis of 53 facial memory and eyewitness identification studies, Deffenbacher, Bornstein, McGorty, and Penrod (2008) found that the negative relationship between time delay and memory is reliable \((r = .18)\). Moreover, Deffenbacher and colleagues suggested the possibility of approximating the probability of a correct identification on a fair six-person lineup.

Retention interval effects are well-illustrated in studies in which recognition accuracy is tested at various time intervals. Results of early research by Egan, Pittner, and Goldstein (1977) indicated that, over time (2 days, 21 days, 56 days), eyewitness participants made substantially
more false identifications. A witnesses’ memory is likely to decline the most within hours or days of the event (Ebbinghaus, 1908). In fact, one study showed that suspect identification rates of witnesses in actual robbery cases dropped the most between 12 hours to 84 hours, with minimal decline thereafter (Tollestrup, Turtle, & Yuille (1994). In another study of actual witnesses, suspects were identified from police lineups more often when the lineup was conducted within one week of the event, compared to when the lineup was conducted in excess of one week of the event (Behrman & Davey, 2001).

In a more recent study, targets stood ten meters from randomly approached individuals for ten seconds, after which participants were asked either immediately (immediate condition) or after an 18-21 day interval (delayed condition) to identify the target from a lineup (Sauer et al., 2009). Accuracy rates for both choosers and non-choosers were lower in the delayed compared to the immediate conditions, though the non-choosers had higher accuracy rates across conditions. Also, the results revealed a relationship between confidence and diagnosticity, such that, across conditions, there was greater diagnosticity among choosers when their confidence was high compared to low (Sauer et al.). Moreover, though confidence levels were higher in the delay compared to the immediate conditions, diagnosticity was lower (Sauer et al.). Thus, while the relationship between confidence and accuracy has received criticism (e.g., Shaw, 1996; Wells, Olsen, & Charman, 2002) there may be circumstances in which eyewitness confidence is a useful indicator of accuracy.

**Eyewitness Identification Reform**

One purpose of research on variables that affect eyewitness identification is arguably to lead to reform (or a reduction in suggestiveness) in identification procedures, such as unbiased lineup instructions and careful attention to avoid giving confirmatory feedback. In Massachusetts,
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

for example, the Supreme Court suggested a set of recommended best practice guidelines that should be used in police training to eliminate suggestive procedures and ensure accurate identifications (Massachusetts Supreme Court, 2013).

One proposed option for reform, as suggested in the lineup instruction literature, is to give witnesses an explicit option not to choose a suspect from a lineup. Indeed, Weber and Perfect (2012) found that witnesses who were given the option to choose “don’t know” in a showup procedure were more accurate and more diagnostic compared to witnesses who were forced to choose “yes” or “no” in the showup procedure. Clark (2012) argued it is important to keep in mind that research efforts made to lower the number of false identifications can also frequently result in a lower number of correct identifications. In an analysis of the utility of eyewitness identification reform, Clark further argued that the cost of a lower number of correct identifications could outweigh the benefit of a lower number of misidentifications.

Are the correct identifications or “hits” we see in these empirical analyses actually reflective of real-life correct identifications? According to Wells, Steblay, and Dysart (2012), the experimental situations in which an eyewitness endures suggestive procedures and then correctly identifies the culprit may not actually be representative of a real-life correct identification. In a brief follow-up to Clark (2012), Newman and Loftus (2012) argue that once an eyewitness is biased by a suggestive procedure, the identification no longer depicts the person’s memory of the event or perpetrator – it simply depicts their memory as a result of suggestion. An accurate correct identification should be based solely upon the eyewitness’ memory, regardless of suggestive procedures (Wells et al.). Clark’s argument further exemplifies the need for eyewitness identification reform, as the only tangible way to prevent eyewitness bias is to remove suggestion from the identification procedures completely.
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Wixted and Mickes (2012) suggest, the real problem lies in the overreliance on the probative value (or the likelihood that an identified suspect is actually guilty) of an identification procedure. Instead, they argued that researchers should utilize receiver operating characteristic (ROC) analyses, which can be used to represent the full range of hit rates and false alarm rates for any given identification procedure, and is therefore the, “only way to convincingly determine which lineup procedure is diagnostically superior” (Wixted & Mickes, p. 277). The suggestions for reform, however, become moot once the unreliable identification enters the courtroom. Thus, in addition to establishing methods to reduce suggestiveness during the identification procedure, it is important to focus on establishing ways in which we can inform jurors about the potential unreliability of an identification.

Established Safeguards Against Unreliable Identifications

Because of the continued concern regarding the reliability of eyewitness evidence, several courts have determined that the judge’s instructions should include mention of some factors thought to affect the reliability of the eyewitness identification, especially because lay knowledge of the influence of these factors is often inadequate (Neal, Christiansen, Bornstein, & Robicheaux, 2012; Desmarais & Read, 2011). Indeed, a specific set of instructions was developed following the appeal of a conviction in which the eyewitness identification was the sole piece of evidence (United States v. Telfaire, 1972). The Telfaire instruction is a widely used, standardized instruction which covers several eyewitness issues, including: viewing time, distance, lighting, prior exposure to the defendant, how the defendant was presented to the witness for identification, and time lapse between crime and identification. Over the years, the Telfaire instruction has garnered the attention of researchers (e.g., Greene, 1988; Cutler, Dexter, & Penrod, 1990). While most of the resulting research suggests that the instructions do not
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

improve jurors’ abilities to accurately evaluate eyewitness identification reliability, there are exceptions. In an early study, jurors rendered significantly fewer guilty verdicts after being presented with the standard judicial instruction, plus a summation of the case, plus judicial commentary regarding eyewitness identification accuracy, compared with jurors who were given either the standard instruction plus judicial summation or the standard instruction alone (Katzev & Wishart, 1985). In the first experiment of a follow-up investigation, Greene showed jurors a videotaped trial in which the quality of the eyewitness identification was manipulated (strong: good lighting, close distance, unobstructed view; weak: bad lighting, further distance, obstructed view). Jurors were randomly assigned to either receive the Telfaire instruction or no eyewitness instruction. In the second experiment, Greene included an additional instruction condition in which the language was simplified and trial specific eyewitness issues were included. If the instruction is an effective method for sensitizing jurors to eyewitness identification accuracy, there should be an increase in guilty verdicts when the identification is strong and a decrease in not guilty verdicts when the identification is weak. Results of the first experiment showed that the Telfaire instruction did not sensitize jurors to the accuracy of the identification. Moreover, in the second experiment, the revised instruction resulted in a skepticism effect, such that jurors rendered more not guilty verdicts regardless of whether the eyewitness identification was strong or weak (Greene).

In another study, jurors were sensitized to the eyewitness evidence after being given a warning by the judge before being introduced to the eyewitness evidence (Ramirez, Zemba, & Geiselman, 1996). On the other hand, Martire and Kemp (2009) found neither sensitivity to eyewitness evidence nor skepticism as a function of judicial instruction. Indeed, a recent meta-analysis showed that neither the Telfaire instruction, nor a revised version of the Telfaire
In related work, Paterson, Anderson, and Kemp (2012) examined whether a judicial warning about relying on co-eyewitness agreement as an indicator of identification accuracy would sensitize jurors to the issues associated with co-witness discussion. To that end, the researchers manipulated whether or not the co-eyewitness had discussed the event details after talking with the police (when the witnesses spoke to one another they were deemed inconsistent – because one of the witnesses altered their statement to match the other witness’ statement post-discussion), as well as whether the judge issued a general or specific warning about the potential effects of co-witness discussion on memory (Paterson et al., 2012). Although there was no evidence of a sensitivity effect for verdicts, jurors generated a greater overall belief in the evidence from the consistent witness compared to the inconsistent witness, and this effect was larger when the judge issued a specific compared to a general warning (Paterson et al., 2012). Moreover, there was no evidence of a skepticism effect (Paterson et al., 2012).

An alternative approach to sensitizing jurors is evident in a study in which researchers tested whether jurors’ ability to evaluate eyewitness identification accuracy could be improved through a teaching aid (Interview-identification-eyewitness; I-I-Eye; Pawlenko, Safer, Wise, & Holfield, 2013). The purpose of this three-step method is to encourage jurors to evaluate whether the eyewitness interview was conducted properly (e.g., could any of the police procedures have contaminated the eyewitness’ memory or confidence in their identification?), to provide guidelines for assessing the eyewitness identification and subsequent interview(s), and to evaluate how variables at the crime scene may have affected the accuracy of the identification (Wise et al., 2009). Jurors viewed one of the following three teaching aids via PowerPoint: the
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Jury Duty aid (basic instruction regarding the defendant’s rights and the importance of being fair and impartial), the Neil v. Biggers aid (description of five eyewitness factors that jurors should consider in their evaluation of the identification), or the I-I-Eye aid (Pawlenko et al., 2013), and subsequently read a 27-page transcript and rendered a verdict. Results indicated a sensitivity effect, such that participants who received the I-I-Eye training rendered more guilty verdicts when the identification was strong compared to when the identification was weak (Pawlenko et al). While the presence of a sensitivity effect is promising, the costly (creation and presentation of PowerPoint slides for each trial) and time-consuming (allowing jurors to go through PowerPoint presentation at own pace) implementation of a teaching aid in the courtroom is not the most legally-plausible approach.

Although existing research suggests that judicial instructions may not be the best way to sensitize jurors to eyewitness identification reliability, courts continue to rule in favor of eyewitness instructions. As discussed earlier, the Supreme Court of New Jersey recently ruled that judges should give jurors tailored instructions regarding any relevant system and estimator variables so that jurors are better equipped to evaluate the reliability of the identification (New Jersey v. Henderson, 2011). The decision to implement these issue-specific instructions was made without the support of any empirical research showing that the instructions effectively sensitize jurors to identification reliability.

Since the implementation of the Henderson instructions in New Jersey courts, researchers have yet to find evidence that the instructions sensitize jurors to eyewitness identification quality. In fact, Yokum and Papailiou (in progress) found only evidence of a skepticism effect, such that conviction rates decreased significantly in the presence of the Henderson instructions (compared to standard judicial instructions) regardless of identification quality. Yokum and Papailiou used
an online sample of participants who watched an abbreviated simulated murder trial in which the evidence was either weak or strong and the instruction at the end of the trial was either standard or enhanced (i.e., a Henderson-type instruction). Perhaps instructions will yield greater sensitivity to identification evidence if they are presented prior to the presentation of the eyewitness evidence. In another study, Laub, Kimbrough, and Bornstein (in progress) found evidence of neither a sensitivity effect nor a skepticism effect using a condensed version of the Henderson-type instruction. A condensed version of the instructions included in a written trial transcript, however, may not yield the same results as the full Henderson instruction read by a judge.

Given the past research showing that instructions of this sort do not generally work as intended (see Berman & Penrod, 2014), it is important to evaluate whether the Henderson instructions will be successful, especially before more states begin to implement similar instructions (see Massachusetts Supreme Court recommendation, 2013; Commonwealth v. Gomes, 2009). Future research should evaluate the efficacy of the Henderson instructions in an ecologically-rich manner (see Diamond, 1997), using adult community member mock jurors.

**Overview**

The current studies sought to examine whether research-based issue-specific judicial instructions (currently being implemented in New Jersey) sensitize jurors to variations in eyewitness identification accuracy. Mock jurors were shown a video trial simulation (Study 1) or a trial transcript (Study 2), in which the judicial instructions, system variables, and estimator variables were manipulated. Following the trial, jurors completed post-trial questionnaires wherein they rendered a verdict and answered questions regarding their perceptions of each of the trial players and the evidence. In Study 1, participants deliberated as a jury until reaching a
unanimous verdict. Once a verdict was reached, participants completed a brief post-deliberation questionnaire. After all participants completed their questionnaires, they were compensated, thanked, debriefed, and dismissed.

**STUDY 1**

The *Henderson* court specified that enhanced instructions are “to be included in the court’s comprehensive jury charge at the close of evidence. In addition, instructions may be given during trial if warranted… Trial courts retain discretion to decide when to offer instructions” (p. 123). As per this recommendation, study 1 sought to examine whether presentation of the *Henderson* instructions prior to the eyewitness testimony will increase sensitivity to identification quality as compared to both the presentation of the *Henderson* instructions at the end of the trial and no judicial instruction on eyewitness testimony. Previous research has shown that judicial instruction on the requirements of proof, presumption of innocence, burden of proof, and reasonable doubt prior to presentation of the evidence, can lower mock jurors’ conviction rates (Kassin & Wrightsman, 1979). In another study, however, mock jurors who received a fact sheet outlining the research related to eyewitness identification prior to the witness testimony were more likely to convict than those who received the fact sheet after hearing the facts of the case (Moore, 2010).

Based on the Court’s assumption in *New Jersey v. Henderson*, that the jury instruction should aid jurors in analyzing and evaluating the credibility of eyewitness identification, we hypothesized that the presence of the *Henderson* instructions would sensitize jurors to the quality of the eyewitness identification. This sensitization will appear in the form of a statistical interaction between witnessing conditions and instructions, such that jurors who hear the
Henderson instructions will give less weight to an identification made under the poor witnessing conditions and more weight to an identification made under the good witnessing conditions (H1). Jurors who do not hear the Henderson instructions will display little or no sensitivity to the manipulated witnessing conditions (i.e. no effects of the manipulated system and estimator variables on verdict or evidence perceptions). The Special Master in New Jersey v. Henderson argued that “whether the science confirms commonsense views or dispels preconceived but not necessarily valid intuitions, it can properly and usefully be considered by both judges and jurors in making their assessments of eyewitness reliability” (p. 124). Based on the Henderson court’s implicit assumption that jurors should be able to evaluate eyewitness identification evidence more effectively after hearing the new instructions, we hypothesized that jurors’ perceptions of the impact of the manipulated system and estimator variables on identification accuracy would mediate the relationship between our independent variables and verdict (H2). Thus, jurors who hear the Henderson instructions will rate the manipulated system and estimator variables as having a greater impact (either positive or negative, dependent upon manipulated evidence quality) on identification accuracy (compared to jurors who do not hear the Henderson instructions), which will then influence jurors’ final verdict decisions.

Method

Participants. Based on the results of a power-analysis, the study required 25 participants per cell to detect a medium to large effect. Three hundred ninety-six jury-eligible community members (33 per cell -- 59% female; 41% male) were recruited via Craigslist.com. Participants came from diverse backgrounds (37.6% White, non-Hispanic; 15.4% Hispanic; 30.8% Black, non-Hispanic; 16.2% Other) and ranged in age from 18-86 ($M = 33.67$, $SD = 13.08$). Participants received $30 in exchange for participating in the three-hour study.
Research design. This study had a 3 (Instruction: No Eyewitness Instruction v. pre-Henderson v. post-Henderson) X 2 (Quality of system variables: Good v. Poor) X 2 (Quality of estimator variables: Good v. Poor) fully-crossed factorial design.

Trial Stimulus. The trial video, New York v. William Thomas Johnson, was between 45-67 minutes (depending upon condition) and was adapted by the researchers from transcripts and videos of a case that was overturned due to a mistaken eyewitness identification (Kentucky v. Gregory, 1993). All names were changed so that the case could not be identified. The trial included opening and closing judicial instructions, opening and closing statements from both attorneys, as well as direct and cross-examination of a police officer, an eyewitness, an expert witness, the defendant, and friend of the defendant. In the trial, the state charged the defendant with attempted rape in the first degree, alleging that the defendant entered the victim’s residence and attacked her in her bed. The expert witness was a forensic serologist who testified that the hair found at the crime scene matched the hair taken from the defendant on 15 out of 15 characteristics and that there is a 1 in 1000 chance that the hair found at the crime scene was not from the defendant. The defendant testified that he was at home sleeping at the time of the crime and his friend testified to that as well.

Estimator Variables. Each of the manipulated estimator variables (i.e., memory decay, weapon focus, and duration) was chosen based on the Henderson court’s decision that variations in these variables may have significant effects on eyewitness identification accuracy and are the subject of new jury instructions. Specifically, the court in Henderson notes that: “delays between the commission of a crime and the time an identification is made can affect reliability” (p. 79); the presence of a weapon can “impair a witness’ ability to make a reliable identification and describe what the culprit looks like” (p. 73); and “the amount of time an eyewitness has to
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

observe an event may affect the reliability of an identification” (p. 74; New Jersey v. Henderson, 2011). The levels of each manipulated estimator variable were based on previous research showing significant differences in identification accuracy between levels. The manipulated estimator variables were the delay between the crime and subsequent identification (e.g., Shapiro & Penrod, 1986; Deffenbacher et al., 2006), weapon focus (e.g., Steblay, 1992; Fawcett, Russell, & Peace, 2013), and exposure duration (e.g., Shapiro & Penrod, 1986; Bornstein et al., 2012). In the good estimator conditions the time delay was one day, there was no weapon present, and the exposure duration was 45 seconds. In the poor estimator conditions the time delay was one month, there was a weapon present, and the exposure duration was 10-15 seconds.

System Variables. Each of the manipulated system variables (i.e., identification procedure, lineup instructions, and confirmatory feedback) were chosen based on the Henderson court’s decision that variations in these variables may have significant effects on eyewitness identification accuracy and are the subject of new jury instructions. Specifically, the court in Henderson notes that: showups, compared to lineups, “fail to provide a safeguard against witnesses with poor memories or those inclined to guess, because every mistaken identification in a showup will point to the suspect” (p. 69); informing the witness that the suspect may or may not be in the lineup is “regarded as one of the most useful techniques for enhancing the reliability of identifications” (p. 53); and confirmatory feedback can “reduce doubt and engender a false sense of confidence” and enhancement in a “witness’ recollection of the quality” of the identification (p. 58; New Jersey v. Henderson, 2011). The levels of each manipulated system variable were based on previous research showing significant differences in identification accuracy between levels. The manipulated system variables were the identification procedure (e.g., Steblay, Dysart, Fulero, & Lindsay, 2003), lineup instructions (e.g., Steblay, 1997;
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Quinlivan, 2012), and confirmatory feedback (e.g., Douglass & Steblay, 2006; Steblay, Wells, & Douglass, 2014). In the good system conditions, the police officer presented a lineup to the witness, informed the witness that the perpetrator may or may not be in the lineup, and did not tell the witness that she chose the suspect. In the poor system conditions, the police officer presented a showup to the witness, failed to mention that the perpetrator may or may not be present, and informed the witness that she chose the suspect.

**Instruction Manipulation.** In all conditions, the judge provided standard instructions at the end of the trial, which included the definition of the charges brought against the defendant, as well as burden of proof and proof beyond a reasonable doubt. These were the only instructions jurors in the no eyewitness instruction condition received. In the pre-Henderson conditions, the judge provided research-based issue-specific instructions prior to the eyewitness (victim) testimony. In the post-Henderson conditions, the judge provided the same research-based issue-specific instructions at the end of the trial (after defense closing statements). As per the *Henderson* court’s recommendation that instruction on system and estimator variables be included only when relevant to a specific case, mention of the manipulated system and estimator variables were only included in the *Henderson* instructions when present in the specific condition (e.g., research on the weapon focus effect was only included in *Henderson* instructions when there was a weapon present in that condition). In all conditions, participants received a written version of the judicial instructions in addition to the verbal presentation (see Appendices for full instructions).

**Measures**

**Voir dire questionnaire.** Prior to watching the trial stimulus, all participants completed a questionnaire regarding their general background and their demographics. Participants provided
information about their gender, major, political views, and ethnic background. To ensure that all participants were jury-eligible and therefore eligible for the study, participants noted whether they were U.S. citizens, if they were registered to vote, if they had a driver’s license, and their age. The questionnaire also asked whether the participants have ever been called for or served jury duty, and if so whether it was for a criminal or civil trial, and whether they were the foreperson on the jury.

**Post-trial measures.** Immediately following the trial, participants completed a post-trial questionnaire assessing verdict as well as evidence and witness perceptions (see Appendix D).

**Verdict.** Participants indicated whether they believed the defendant was guilty or not guilty of attempted rape, their confidence in their verdict (1 = not at all confident; 9 = extremely confident), and the probability that the defendant was guilty (0% = not at all probable; 100% = completely probable).

**Evidence perceptions.** Participants indicated, on a scale of 1 to 9 (where 1 represented strongly reduced accuracy and 9 represented strongly increased accuracy), the impact of each of the manipulated variables (lineup instructions, duration, weapon, memory decay, identification procedure, confirmatory feedback) on the likelihood of an accurate identification.

**Witness ratings.** Participants indicated, on a scale of 1 to 7 their impressions of each of the witnesses’ trustworthiness (1 - trustworthy; 7 - untrustworthy), honesty (1 – honest; 7 – dishonest), convincingness (1 – convincing; 7 – unconvincing), and certainty (1 – certain; 7 – uncertain; Detective: $\alpha = .87$, Victim: $\alpha = .85$, Defendant: $\alpha = .86$).

**Deliberation.** Immediately following the post-trial questionnaire, participants were instructed to gather around a table to deliberate as a jury. All deliberations were videotaped and jurors were given a maximum of 45 minutes to deliberate. If juries were not able to reach a
verdict after 40 minutes they were urged to do their best to reach a verdict and allowed an extra five minutes to do so. Once juries reached a verdict, each juror completed a post-deliberation verdict form and the appointed foreperson completed a group jury verdict form. Results of the deliberations are included in a separate analysis (see Bergold, Jones, Berman, & Penrod, 2014).

Procedure

Prior to the beginning of the study, participants provided informed consent and completed a voir dire questionnaire. Once the informed consent and voir dire questionnaires were collected, participants watched the simulated trial. After the trial concluded, participants completed a post-trial questionnaire and subsequently began deliberation. Once all participants completed the post-deliberation verdict forms they were debriefed, compensated, and dismissed.

Pilot Study

A pilot study was conducted to examine what we believed would lead to the highest and lowest conviction rates, in order to determine whether there were any ceiling or floor effects. Participants included 278 undergraduates, 30 community members, and 67 mTurkers (N = 375). The system variables manipulated were: multiple viewings (photo lineup + showup v. photo lineup only); lineup instructions (biased v. unbiased); suspect presentation (showup v. lineup); and confirmatory feedback (present v. absent). The estimator variables manipulated were: time lapse (one day v. one week); disguise (present v. absent); viewing time (15 sec v. 45 sec); and weapon focus (knife v. no knife).

In the low conviction condition, all system and estimator variables were bad (e.g., bad system: biased instructions; bad estimator: lesser viewing time) and there were both issue-specific eyewitness judicial instructions and expert testimony. In the high conviction condition,
all system and estimator variables were good (e.g., good system: unbiased instruction; good estimator: longer viewing time) and there was no eyewitness instruction.

Results of the initial two-cell experiment ($n = 40$) indicated that there were not enough guilty verdicts in the high conviction condition, so we made several modifications to the high conviction condition in an effort to increase the number of convictions and still ensure that the general weight of the trial evidence was ambiguous. These modifications included manipulations of: the amount of details in the victim’s description of the perpetrator (little v. many), victim certainty (hesitant v. not hesitant in identification), mention of the victim’s scratches from the perpetrator (present v. absent), the presence of the defendant (present v. absent), the presence of a forensic serologist (present v. absent), mention of the police “egging on” the defendant on the way to the police station (present v. absent), the inclusion of the burglary charge in addition to the attempted rape charge (present v. absent), and the combined number of system and estimator variables included (six v. eight). System and estimator variables were first manipulated individually, and later grouped together once we realized that they did not result in any differences in verdicts independently.

Results indicated that when the high conviction condition included: little detail in the victim’s description of the perpetrator, no hesitation in identification, no mention of scratches from the perpetrator, presence of the defendant, presence of the forensic serologist, no mention of police “egging on” defendant, no inclusion of burglary charge, and six combined system and estimator variables, there was an acceptably larger number of guilty verdicts in the high conviction condition (50%) compared to the low conviction condition (25%, $n = 209$). This difference in verdict was determined to be sufficient to detect differences in verdicts between conditions, should they exist. Thus, this version of the case was used in the proposed study.
Results

Results are presented mostly in the form of a hierarchical path model (generated using SPSS software) wherein the first block (Step 1) represents the manipulated variables, the second block (Step 2) represents the interactions between the manipulated variables, the third block (Step 3) represents the hypothesized mediating variables, and verdict is the final dependent variable.

Manipulation checks

Significantly more people reported that the judge specifically mentioned all three of the manipulated estimator variables (i.e., memory decay, weapon focus, and duration) when they heard Henderson instructions (pre: 63.0%; post: 65.1%) compared to those that did not hear Henderson instructions (2.5%; $\chi^2 = 181.00, p < .001$). Additionally, significantly more people reported that the judge specifically mentioned all three of the manipulated system variables (i.e., lineup instructions, confirmatory feedback, and showup) when they heard the instructions (pre: 44.6%; post: 48.1%) compared to those that did not hear the instructions (3.9%; $\chi^2 = 138.68, p < .001$). The majority of participants correctly reported the presence or absence of the manipulated system variables (95.5% got at least 2 of 3 questions correct) and estimator variables (82.5% got at least 2 of 3 questions correct). There were no differences in the pattern of results as a function of whether participants answered manipulation checks correctly, thus all participants were included in the analyses.

Variables included in analyses

Step 1 path analysis variables. Four variables were included in Step 1. Two variables were two-level independent variables (System variables, Estimator variables) and the other two
were dummy variables (pre-\textit{Henderson}, post-\textit{Henderson}) representing comparisons between instruction conditions with No Eyewitness Instruction as the baseline.

\textbf{Step 2 path analysis variables.} Nine variables were included in Step 2. In addition to the variables in Step 1, variables in Step 2 included the interactions between independent variables (pre-\textit{Henderson} X system, pre-\textit{Henderson} X estimator, post-\textit{Henderson} X system, post-\textit{Henderson} X estimator, system X estimator).

\textbf{Individual mediator analyses.} Nine individual regressions were run on variables that we hypothesized may mediate the relationship between our independent variables and verdict. Six variables concerned perceptions of evidence (lineup instructions, duration, weapon, memory decay, procedure, and confirmatory feedback). The remaining three variables refer to jurors’ ratings of the witnesses (detective, defendant, and victim).

\textbf{Step 3 path analysis variables.} Eighteen variables were included in Step 3. In addition to the variables in Step 1 and Step 2, variables in Step 3 included all hypothesized mediator variables (evidence perceptions and witness ratings).

\textbf{Verdict}

We created a continuous guilt scale for confidence-weighted verdict, by multiplying participants’ confidence by +1 if they rendered a guilty verdict and by -1 if they rendered a not guilty verdict. This measure, however, was bimodal, so we chose to use the dichotomous verdict measure for all subsequent analyses.

The overall model accounted for a substantial amount of the variance in jurors’ verdicts ($R^2 = .60$, $F(20, 435) = 35.20$, $p < .001$). All total, direct, and indirect effects are presented in Table 2, and all significant direct paths are presented in Table 3. For all analyses, positive $\beta$ indicates increase in convictions; negative $\beta$ indicates decrease in convictions.
(H1) Did Henderson instructions increase sensitivity to eyewitnessing conditions?

Results of Step 2 indicated that the Henderson instructions did not produce the hypothesized sensitivity effect. Specifically, the interactions between the Henderson instructions and eyewitnessing conditions did not significantly affect participants’ verdicts (all ps > .21; $R^2 = .02, F(9, 446) = 2.24, p = .02$). Overall, those who received either Henderson-before (32.2%) or Henderson-after (39.8%) instruction rendered fewer guilty verdicts compared to those who heard no eyewitness instruction (51.6%).

**How did manipulated variables affect verdict?** Results of Step 1 indicated that each of the independent variables (i.e., system variables: $\beta = .09, p = .06$; estimator variables: $\beta = -.10, p = .03$, Henderson-before instructions (versus no eyewitness instructions): $\beta = -.18, p = .001$; Henderson-after instructions (versus no eyewitness instructions): $\beta = -.12, p = .02$) significantly affected verdict ($R^2 = .03, F(4, 451) = 4.80, p = .001$).

**How did manipulated variables affect perceptions of evidence and witnesses?**

**System variables.** The manipulated system variables had direct effects on participants’ ratings of the impact (on identification accuracy) of lineup instructions ($\beta = .51$), weapon presence ($\beta = .14$), memory decay ($\beta = .09$), and identification procedure ($\beta = .33$); as well as an effect of feedback on witness confidence ($\beta = -.56$; all $ps < .05$). System variables also had direct effects on participants’ ratings of the detective ($\beta = .17$), victim ($\beta = .13$) and the defendant ($\beta = -.11$; all $ps < .02$; See Table 4 for descriptive statistics).

**Estimator variables.** Estimator variables had direct effects on participants’ ratings of the impact (on identification accuracy) of lineup instructions ($\beta = -.10$), duration ($\beta = .15$), weapon presence ($\beta = .16$) memory decay ($\beta = .31$) and identification procedure ($\beta = -.11$) on
identification accuracy (all $ps < .03$). Estimator variables also had direct effects on participants’ ratings of the detective ($\beta = -.13$), victim ($\beta = -.10$) and the defendant ($\beta = .13$; all $ps < .03$).

**Henderson-before instructions.** Henderson-before instructions had direct effects on participants’ ratings of the impact (on identification accuracy) of lineup instructions ($\beta = -.10$), duration ($\beta = -.13$), weapon presence ($\beta = -.12$), memory decay ($\beta = -.10$), and identification procedure ($\beta = -.17$; all $ps < .05$).

**Henderson-after instructions.** Henderson-after instructions had direct effects on participants’ ratings of the impact (on identification accuracy) of lineup instructions ($\beta = -.11$), memory decay ($\beta = -.10$), and identification procedure ($\beta = -.13$; all $ps \leq .05$; see Table 4 for all values).

**System variables X estimator variables.** The direct effects of system and estimator variables on ratings of the impact of the procedure on identification accuracy were qualified by a significant interaction ($\beta = -.19, p = .02$), such that when system variables were poor, jurors ratings of the impact of the procedure did not differ as a function of whether estimator variable quality was good ($M = 4.31$) or poor ($M = 4.29; d < .01$). When system variables were good, however, jurors’ ratings of the impact of the procedure on the likelihood of an accurate identification were greater when estimator variables were poor ($M = 6.16$) compared to good ($M = 5.24; d = .31$).

The direct effects of system and estimator variables on jurors’ ratings of the impact of lineup instructions on the likelihood of an accurate identification were qualified by a significant interaction ($\beta = -.17, p = .03$). Specifically, when system variables were poor, jurors’ ratings of the impact of the lineup instructions did not differ as a function of whether estimator variable quality was good ($M = 3.88$) or poor ($M = 3.90; d < .01$). When system variables were good,
however, jurors’ ratings of the impact of the lineup instructions on the likelihood of an accurate identification were greater when estimator variables were poor ($M = 6.59$) compared to poor ($M = 5.76; d = .30$).

**Estimator variables X Henderson-before instructions.** The direct effects of estimator variables and Henderson-before instructions on the impact of weapon presence were qualified by a significant interaction ($\beta = -.23, p = .01$), such that when estimators were poor, jurors rated the weapon as having a more positive effect on identification accuracy (strongly increased accuracy) when they heard no eyewitness instruction ($M = 5.20$) compared to when they heard Henderson-before instructions ($M = 4.02; R^2 = .06, F(9, 458) = 4.49, p < .01$). When estimators were good, there was no difference in jurors’ ratings of the impact of the weapon as a function of instruction type (all $p$s >.65).

(H2) **Did ratings of the evidence and witnesses mediate the relationship between independent variables and verdict?**

We hypothesized that ratings of the evidence and witnesses would serve as mediators between the independent variables and verdict. For the purposes of our study, a mediated relationship occurred when there was a total effect of an independent variable on verdict (significance in Step 1), and when that independent variable had an effect on the mediator (significance in individual regression analyses) which also had an effect on verdict (significance in the Step 3; see Baron & Kenny, 1986). All regression analyses were also computed using the Hayes PROCESS macro (model 10; Hayes, 2013). No difference in results was found between the two methods.

Results indicated that the direct effects of system variables, estimator variables, and Henderson-after instructions on verdict were entirely mediated through ratings of the evidence
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

and witnesses ($R^2 = .53$, $F(20, 435) = 26.21; p < .001$), thus partially supporting our hypothesis. That is, the direct effects became non-significant when taking into account the mediating variables. Specifically, the direct effect of system variables on verdict was mediated by participants’ ratings of the identification procedure ($\beta = .15, p < .01$), the detective ($\beta = .11, p = .01$), and the victim ($\beta = .19, p < .001$). The direct effect of estimator variables on verdict was mediated by participants’ ratings of the duration ($\beta = .13, p = .01$), identification procedure ($\beta = .15, p < .01$), the detective ($\beta = .11, p = .01$), and the victim ($\beta = .19, p < .001$). The direct effect of Henderson-after instructions on verdict was mediated by participants’ ratings of the identification procedure ($\beta = .15, p < .01$) and the detective ($\beta = .11, p = .01$). The direct effect of Henderson-before instructions was only partially mediated by participants’ ratings of the evidence and witnesses ($\beta = -.16, p = .01$).

**Discussion**

Although the Henderson Court reasoned that these research-based issue-specific instructions would help make jurors aware of the potential effects of system and estimator variables on eyewitness identification quality, the results of the current study suggest they are not as effective as anticipated.

**Sensitizing Effects**

If the Henderson instructions effectively sensitized jurors to identification quality (H2), there would have been interactions between instructions (pre- and/or post-) and system and estimator variables on verdict. That is, participants who heard the Henderson instructions would, according to the Henderson Court’s reasoning that these instructions would help jurors evaluate evidence “critically and objectively” (p. 124), render more guilty verdicts for the strong case (good estimator and system variables) and fewer guilty verdicts for the weak case (poor
estimator and system variables) than those who heard no eyewitness instructions. Rather than sensitizing participants, the Henderson instructions (both pre- and post-) led to an overall skepticism effect among participants. The presence of this skepticism effect indicates that the instructions might have produced an overall more critical view of eyewitness evidence or confusion, and may thus be counter to the Henderson court’s hypothesis that the instructions would help to “avoid possible confusion to jurors created by dueling experts” (p. 126). While it is not clear that the dueling expert conjecture is correct (Levett & Kovera, 2008; Levett & Kovera, 2009), the Henderson instructions did not have their intended sensitizing effect. Participants who heard either the Henderson-before (32.2%) or Henderson-after (39.8%) instruction rendered fewer guilty verdicts than those who heard no eyewitness instruction (51.6%), regardless of the quality of the witnessing and identification conditions. Thus, although our data does not support the Henderson court’s assumption that the instructions will help jurors analyze and evaluate the credibility of eyewitness identification, it does indicate that the instructions lead jurors to be more critical of eyewitness evidence.

Although there was no evidence of a sensitivity effect, the manipulated variables did influence evidence perceptions, which is consistent with the Henderson court’s reasoning that “many jurors, through their life experiences and intuition, generally understand how memory works” (p. 123). Specifically, there were many effects of the manipulated variables on participants’ ratings of the evidence and witnesses, most of which reflect a more skeptical view of eyewitness evidence. Thus, our hypothesis that the relationships between our manipulated variables and verdict would be mediated by participants’ ratings of the witnesses and evidence perceptions (H2) was partially supported.

**Impact of IVs on evidence perceptions**
System variables. The manipulated system variables had direct effects on participants’ ratings of the impact of lineup instruction ($\beta = .51$), weapon presence ($\beta = .14$), memory decay ($\beta = .09$), identification procedure ($\beta = .33$), feedback ($\beta = -.56$), and ratings of the detective ($\beta = .17$), victim ($\beta = .13$), and defendant ($\beta = -.11$). All of these effects were in the expected direction, such that increased identification accuracy was associated with good system and estimator variables. Specifically, when system variables were good, participants rated unbiased lineup instructions, absence of a weapon, a short time delay, and use of a lineup (vs. showup) as increasing identification accuracy compared to when system variables were poor. Participants rated confirmatory feedback as increasing the witness’ confidence when the system variables were poor compared to when system variables were good. Additionally, when system variables were good, participants gave more favorable ratings of the detective and the victim, and less favorable ratings of the defendant compared to when system variables were poor.

Estimator variables. The manipulated estimator variables had direct effects on participants’ ratings of the impact of lineup instruction ($\beta = -.10$), duration ($\beta = .15$), weapon presence ($\beta = .16$), memory decay ($\beta = .31$), identification procedure ($\beta = -.11$), and ratings of the detective ($\beta = -.13$), victim ($\beta = -.10$), and defendant ($\beta = .13$). In this case, all effects were in the expected direction with the exception of lineup instructions and witness ratings. Specifically, when estimator variables were good, participants gave higher accuracy ratings for unbiased lineup instructions, a long duration, absence of a weapon, a short time delay, and use of a lineup (vs. showup), as compared to when estimator variables were poor. Surprisingly, when estimator variables were good, participants rated confirmatory feedback as having increased witness confidence compared to when estimators were poor. When estimator variables were good, participants gave less favorable ratings of the detective and the victim, and more favorable
ratings of the defendant compared to when estimator variables were poor. These results suggest that jurors’ perceptions of the evidence quality also influence their overall perceptions of the witnesses.

Instructions. The Henderson-before instructions had direct effects on participants’ ratings of the impact of lineup instructions, duration, weapon presence, memory decay, and identification procedure. These effects indicate a skepticism effect, such that when participants heard Henderson-before instructions, compared to no eyewitness instruction, they rated biased lineup instructions, a short duration, presence of weapon, a long time delay, and showup (vs. lineup) as more strongly reducing the accuracy of the identification, regardless of the quality of system and estimator variables present in the trial.

The Henderson-before instructions interacted with estimator variables to influence ratings of the impact of the weapon on the likelihood of an accurate identification. This interaction, however, was counter to the Henderson court’s prediction that the instructions would assist jurors in objectively and critically evaluating eyewitness identification evidence. If the instructions did effectively sensitize jurors when there was a weapon present, participants should have rated the impact of the weapon as greater when they heard the Henderson instructions compared to when they did not hear the instructions. We found the opposite effect, such that jurors rated the impact of the weapon as lower when they heard the Henderson instructions compared to when they heard no eyewitness instruction. This was not entirely surprising, as research shows that weapon focus is a variable with which most laypeople are not familiar (Desmarais & Read, 2010). In fact, this likely indicates that jurors either misinterpret (i.e. interpret instruction as an indication that the presence of a weapon might enhance eyewitness identification) or have difficulty comprehending the effect of weapon focus after hearing the
Henderson instruction. If this is the case, future studies should focus on simplifying that portion of the instruction to make it more comprehensible for jurors. The effect of Henderson-before instructions remained significant after taking all variables into account, which indicates that, although these instructions influenced participants’ ratings of the evidence and witnesses, the relationship between Henderson-before instructions and verdict was not accounted for by our hypothesized mediators.

The direct effects of the Henderson-after instruction compared to no eyewitness instruction, on participants’ ratings of the impact of lineup instructions, memory decay, and identification procedure further supports the presence of a skepticism effect. The effect of Henderson-after instructions on verdict was completely mediated by participants’ ratings of the identification procedure. Specifically, ratings of the impact of the identification procedure on the likelihood of an accurate identification were greater when participants heard Henderson-after (increased accuracy) compared to no eyewitness instructions (decreased accuracy). Participants’ ratings of the identification procedure were then related to verdict, such that when the procedure was perceived to increase accuracy guilty verdicts also increased.

**Evidence Mediation**

**System variables --> verdict.** The direct effect of system variables on verdict was completely mediated by participants’ ratings of the identification procedure, detective, and victim. That is, the direct effect of system variables on verdict became non-significant after taking into account the mediating variables. Specifically, ratings of the impact of the identification procedure on the likelihood of an accurate identification were greater when the witness viewed a lineup (increased accuracy) compared to a showup (decreased accuracy) ($\beta = .33$). Ratings of the identification procedure were related to verdict, such that greater impact
ratings were associated with more guilty verdicts ($\beta = .16$). Ratings of the detective ($\beta = .17$) and the victim ($\beta = .13$) were greater when system variables were good compared to poor. These ratings were then related to verdict, such that greater ratings of the detective ($\beta = .11$) and the victim ($\beta = .22$) resulted in more guilty verdicts.

**Estimator variables -> verdict.** The direct effect of estimator variables on verdict was completely mediated by participants’ ratings of duration, identification procedure, detective, and victim. That is, the direct effect of estimator variables on verdict became non-significant after taking into account the mediating variables. Specifically, participants rated the longer duration, as having a greater impact on the likelihood of an accurate identification, compared to shorter duration ($\beta = .15$). Ratings of the impact of the duration were then related to verdict, such that higher impact ratings were related to more guilty verdicts ($\beta = .13$). Ratings of the detective ($\beta = -.13$) and victim ($\beta = -.10$) were lower when estimator variables were poor compared to good. Participants’ ratings of the detective and victim were then related to verdict, such that lower ratings of the detective ($\beta = .11$) and the victim ($\beta = .22$) were related to more acquittals.

The complete mediation of system and estimator variables on verdict may indicate that jurors do not need judicial instruction to understand the impact of some of these variables. That is, paths from the mediators to verdict indicate that jurors’ impressions of the impact of some variables (system: identification procedure; estimator: duration) align with their verdict judgments in a sensible way. The skepticism effects, however, indicate that while jurors’ impressions can be influenced by the manipulated variables, instructions improve neither the quality of evidence evaluations nor jurors’ final judgments. While the *Henderson* court did not explicitly hypothesize skepticism effects, evidence that *Henderson* instructions reduced conviction rates may result in a reduction in erroneous convictions (although at the cost of fewer
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

convictions in cases that may warrant convictions), which was the intention of the Henderson court.

As is the case with most jury decision-making studies, the current study has its limitations. It should be noted, however, that the use of adult community members, stimulus materials based on an actual case, and a video trial renders the current study the most ecologically strong test of the Henderson instruction to date. Overall, research on the efficacy of the Henderson instruction (e.g., Yokum & Papailiou, 2014; Laub, Kimbrough, & Bornstein, 2013) indicates that these instructions are no more effective than previous versions of eyewitness identification jury instructions (i.e., Telfaire).

STUDY 2

Overview

The purpose of the second study was to determine whether modifications to the existing Henderson instructions would increase jurors’ sensitivity to identification quality. Specifically, this study investigated whether encouraging jurors to evaluate the eyewitness identification evidence and judicial instruction, prior to rendering a verdict would increase jurors’ sensitization to the identification quality. This hypothesis stems from basic testing effect literature indicating that practice-testing during learning leads to an increase in one’s memory for the information (e.g., Gates, 1917; Carrier & Pashler, 1992; Roediger & Karpicke, 2006a). Thus, answering questions about the evidence and the potential impact of the manipulated variables on the identification (i.e., retrieval practice) should enhance jurors’ memory and hopefully their understanding and interpretation of the instructions.

The Henderson instructions were modified to be more succinct (incorporating prompts, described below, used in Pawlenko et al., 2013) and were manipulated such that jurors heard
either no eyewitness instruction, the case-specific Henderson instruction, or the all-inclusive Henderson instruction (see more detailed description of instruction manipulation below). Moreover, we utilized a different participant pool (Amazon’s Mechanical Turk) and trial stimulus and manipulated different system and estimator variables, to increase the generalizability of the results.

Hypotheses

H1: The presence of all-inclusive Henderson instructions will result in greater sensitivity to identification quality compared to case-specific Henderson instructions and No eyewitness instructions (because the Henderson instructions are drawing attention to the poor as well as the good eyewitness identification evidence).

H2: Evaluation of evidence before rendering a verdict (rather than after) will increase sensitivity to identification quality.

Method

Participants. Based on the results of a power analysis, the study required twenty-five participants per cell to detect a medium to large effect. Thus, two hundred and ninety-five jury-eligible community member participants (24.6 per cell -- 45% female; 55% male) were recruited via Amazon’s Mechanical Turk (mTurk). Participants ranged in age from 18 to 71 ($M = 36$) and had diverse backgrounds (73% White, non-Hispanic; 6% Hispanic; 9% Black, non-Hispanic; 12% Other). In exchange for their participation, participants received $1.50.

Research design. This study had a 3 (Instruction: Henderson case-specific v. Henderson all-inclusive v. No Eyewitness Instruction) X 2 (System/Estimators: Good v. Poor) X 2 (Evidence Evaluation: Pre-verdict v. Post-verdict) fully-crossed factorial design.
**Trial Stimulus.** The trial summary, *New York v. Christopher Bryan*, was approximately four pages long (≈1,670 words). The trial included summaries of opening and closing judicial instructions, opening and closing statements from both attorneys, as well as summaries of the direct and cross-examination of a police officer, the victim (eyewitness), and the defendant’s girlfriend. In the trial, the state charged the defendant with armed robbery, alleging that the defendant entered the victim’s jewelry store and held her at gunpoint as she gave him the money in the safe. As the trial stimulus was novel, the pilot test compared two conditions in which jurors were expected to render either more guilty or more not guilty verdicts to ensure an adequate verdict split (57% guilty in guilty condition; 36% guilty in not guilty condition).

**Estimator Variables.** The manipulated estimator variables were lighting and exposure duration. In the good estimator variables conditions the lights were turned on in the jewelry store and the witness saw the perpetrator for one minute. In the bad estimator variable conditions, the lights were turned off in the jewelry store and the witness saw the perpetrator for less than 10 seconds.

**System Variables.** The manipulated system variables were multiple viewings and lineup instructions. In the good system variables conditions, the witness saw only a lineup in which the suspect was present and the administrator told the witness that the suspect may or may not be in the lineup. In the bad system variables conditions, the witness saw both a set of mugshots and a lineup in which the suspect was present (no identification was made from the mugshots) and the administrator did not tell the witness that the suspect may or may not be in the lineup.

**Henderson Instruction Manipulation.** In the *Henderson* case-specific conditions, the judge provided abbreviated research-based issue-specific instructions prior to eyewitness testimony; mention of the manipulated system and estimator variables were only included in the
NJ instruction when present in the specific condition. In the Henderson all-inclusive conditions, the judge provided abbreviated research-based issue-specific instructions prior to eyewitness testimony; mention of the manipulated system and estimator variables were included in the NJ instruction regardless of whether system and estimator variables were good or poor. For both Henderson instructions, the judge prompted jurors to think about the potential impact of each variable on the identification (e.g., after description of duration, “Ask yourself: For how long did the eyewitness view the perpetrator? Did this enhance or impair her accuracy?” see Appendix E). In the No Eyewitness Instruction condition, the judge provided standard instruction on burden of proof and proof beyond a reasonable doubt (see Figure 1).

Evidence Evaluation Manipulation. In the Pre-verdict condition, jurors were asked to evaluate the evidence before rendering a verdict. In the Post-verdict evaluation condition, jurors were asked to render a verdict immediately after the trial concluded.

Measures

Voir dire questionnaire. Prior to reading the trial transcript, all participants answered questions regarding their general background and questions about themselves. Participants provided information about their gender, political views, and ethnic background. To ensure that all participants were eligible for the study, they reported whether they were U.S. citizens, if they were registered to vote, if they have a driver’s license, and their age. Participants also reported whether they had ever been called for or served jury duty, and if so whether it was for a criminal or civil trial, and whether they were the foreperson on the jury.

Post-trial measures. Immediately following the trial, participants completed a post-trial questionnaire assessing verdict as well as evidence and witness perceptions.
**Verdict.** Participants indicated whether they believed the defendant was guilty or not guilty of armed robbery, their confidence in their verdict (1 = not at all confident; 9 = extremely confident), and the probability that the defendant was guilty (0% = not at all probable; 100% = completely probable).

*Planned comparisons of interest.* Several variables were computed to further test the effects of the manipulated variables and their interactions on verdict. Based on our hypotheses, there should be a two-way interaction of instruction type and evidence quality on verdict as well as a three-way interaction of instruction type, evidence quality, and evidence evaluation on verdict. We would expect, for example, to see a significant difference in verdict between those in the condition in which there was no eyewitness instruction, evidence evaluation before verdict, and poor evidence quality compared to the condition in which there was no eyewitness instruction, evidence evaluation before verdict, and *good* evidence quality. To determine whether such a difference was significant, we selected the data from all relevant conditions and created a dummy variable in which the former condition was coded as a 1 and the latter condition was coded as a -1. Each dummy variable was then individually included in the overall model.

*Comparison 1.* No eyewitness instruction, evidence evaluation after verdict, good evidence quality vs. no eyewitness instruction, evidence evaluation after verdict, poor evidence quality.

*Comparison 2.* No eyewitness instruction, evidence evaluation before verdict, good evidence quality vs. no eyewitness instruction, evidence evaluation before verdict, poor evidence quality.


Comparison 5. All-inclusive Henderson, evidence evaluation after verdict, good evidence quality vs. All-inclusive Henderson, evidence evaluation after verdict, poor evidence quality.

Comparison 6. All-inclusive Henderson, evidence evaluation before verdict, good evidence quality vs. All-inclusive Henderson, evidence evaluation before verdict, poor evidence quality.


Comparison 8. All-inclusive Henderson, evidence evaluation before verdict, poor evidence quality vs. No eyewitness instruction, evidence evaluation before verdict, poor evidence quality.


Comparison 12. All-inclusive Henderson, good evidence quality vs. All-inclusive Henderson, poor evidence quality.

Comparison 13. All-inclusive Henderson, poor evidence quality vs. No eyewitness instruction, poor evidence quality.


Evidence perceptions. Participants indicated, on a scale of 1 to 9 (where 1 represented strongly reduced accuracy and 9 represented strongly increased accuracy), the impact of each of the manipulated variables (multiple viewings, lineup instructions, lighting, exposure duration) on the likelihood of an accurate identification.

If/Then Statements. This set of questions was designed to further gauge participants’ knowledge and decision-making abilities regarding the impact of the manipulated variables. Kahneman and Tversky (1981) argue that, when considering an alternative scenario, people more readily imagine an alternative scenario that is normal versus exceptional. After hearing the Henderson instructions, participants in the poor eyewitnessing conditions should thus be able to imagine a situation in which the eyewitnessing conditions were good, and reach a hypothetical verdict accordingly. Participants indicated whether alterations in the evidence quality of the manipulated variables would have an effect on their final verdicts (e.g., “Would your final verdict be different if the lighting had been better at the time of the crime?” and “Would your final verdict be different if the police had told the witness that the suspect may or may not be in the lineup?”). These questions were based on Charman and Wells (2008) actual/counterfactual
paradigm, wherein witnesses were asked to imagine how they would have responded if they had been in a different condition.

Procedure

Participants agreed to participate through Amazon’s Mechanical Turk (mTurk). Participants provided informed consent, completed a voir dire questionnaire, and read the trial transcript. Because of the risks associated with conducting research online, we included a safeguard so that we would be able to determine whether or not participants actually read the judicial instructions. Embedded in the Henderson instructions was an instruction to enter the following sentence when prompted on a later page: “I have read the judge’s instructions.” Upon completion of the trial, participants completed a post-trial questionnaire. Once participants completed the post-trial questionnaires, they were debriefed online and compensated via mTurk.

Results

Results of the effects of the manipulated variables on verdict were calculated using logistic regression analyses. We first analyzed the effect of the overall model (Henderson all-inclusive vs. no eyewitness instruction, Henderson case-specific vs. no eyewitness instruction, evidence quality, evidence evaluation and their interactions on verdict). Next, we ran separate regression analyses wherein each of the planned comparisons were included. Univariate ANOVAs were used to calculate the main effects of the independent variables, as well as their interactions, on participants’ perceptions of the evidence.

Manipulation checks

Significantly more people reported that the judge specifically mentioned three of the four manipulated system and estimator variables (i.e., lineup instructions, viewing time, and lighting) when they heard eyewitness instructions (case-specific Henderson instructions: 32%; non-case-
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

specific Henderson instructions: 63%) compared to those that heard no eyewitness instructions (5%; \( \chi^2 = 136.70, p < .001 \)). The majority of participants correctly reported the presence or absence of the manipulated system variables in the good conditions (86% got all four questions correct) and poor conditions (82% got all four questions correct). There were no differences in the main pattern of results as a function of whether participants answered manipulation checks correctly, thus all participants were included in the analyses.

The pattern of results did not differ as a function of those who reported the embedded sentence regarding reading the judge’s instructions \((N = 136)\) versus those that did not \((N = 58)\). Thus, all participants were included in the analyses.

Sensitivity Effects

The overall 3x2x2 model was significant \((\chi^2 (4) = 21.29, p < .001)\). All effects are presented in Table 5. For all analyses, positive \(\beta\) indicates increase in convictions; negative \(\beta\) indicates decrease in convictions. The results of the regression analysis revealed a significant effect of evidence quality on verdict \((\beta = .92, \text{Wald} = 5.10, \text{Exp} (B) = 2.50, p = .02)\), such that there was an increase in convictions when evidence quality was good (62%) and a decrease in convictions when evidence quality was poor (38%). No other main effects or interactions were significant (all \(ps > .14\)).

Planned comparisons of interest. Results indicated that four of fourteen planned comparisons were statistically significant (see Tables 6 and 7 for all values).

Comparison 1: When participants received no eyewitness instruction and evaluated the evidence after rendering a verdict, there were significantly more convictions when evidence quality was good (68%) compared to when evidence quality was poor (41%; \(\beta = 1.13, \text{Wald} = 3.78, \text{Exp} (B) = 3.09, \text{Nagelkerke} R^2 = .10, p = .05\)).
Comparison 3: When participants received case-specific Henderson instructions and evaluated the evidence after the verdict, there were significantly more convictions when evidence quality was good (62%) compared to when evidence quality was poor (28%; $\beta = 1.41$, Wald = 5.54, Exp (B) = 4.11, Nagelkerke $R^2 = .15$, $p = .02$).

Comparison 6: When participants received all-inclusive Henderson instructions and evaluated the evidence prior to rendering a verdict, there were significantly more convictions when evidence quality was good (67%) compared to when evidence quality was poor (20%; $\beta = 2.08$, Wald = 8.65, Exp (B) = 8.00, Nagelkerke $R^2 = .28$, $p = .003$) and this was the largest spread in means among all comparison tests.

Comparison 12: When participants received all-inclusive Henderson instructions, there were significantly more convictions when evidence quality was good (60%) compared to when evidence quality was poor (25%; $\beta = 1.49$, Wald = 10.41, Exp (B) = 4.42, Nagelkerke $R^2 = .16$, $p = .001$).

Evidence perceptions.

Identification procedure. There was a significant main effect of quality of system and estimator variables on ratings of the impact of lineup procedure on the likelihood of an accurate identification ($F (1, 283) = 285.46, p < .001$, $\eta^2 = .21$). Mock jurors rated the lineup procedure as increasing the likelihood of an accurate identification when the quality of the system and estimator variables was good ($M = 6.50$) compared to when the quality of the system and estimator variables was poor ($M = 4.52$).

There was a significant main effect of evidence evaluation on ratings of the impact of lineup procedure on the likelihood of an accurate identification ($F (1, 283) = 3.78, p = .05$, $\eta^2 = .01$). Mock jurors who evaluated the evidence after rendering a verdict were more inclined to
rate the lineup procedure as increasing the likelihood of an accurate identification \( (M = 5.74) \) compared to those who evaluated the evidence prior to rendering a verdict \( (M = 5.29) \).

There was a marginally significant three-way interaction between instruction, evidence quality, and evidence evaluation on ratings of the impact of lineup procedure \( (F(2, 283) = 2.74, p = .07, \eta^2 = .02) \). Specifically, when evidence quality was poor and jurors rendered a verdict after evaluating the evidence, jurors were more inclined to rate the lineup procedure as increasing the likelihood of an accurate identification when they heard no eyewitness instruction \( (M = 5.12) \) compared to when they heard all-inclusive Henderson instructions \( (M = 3.76; F(2, 283) = 3.06, p < .05, \eta^2 = .02) \). No other significant differences as a function of instructions were found when quality was poor and jurors rendered a verdict after evaluating the evidence \( (all \, ps > .14) \). When evidence quality was poor and jurors rendered a verdict prior to evaluating the evidence, there were no significant differences in ratings as a function of instructions \( (all \, ps > .31) \). When evidence quality was good and jurors rendered a verdict after evaluating the evidence, jurors rated the lineup procedure as increasing the likelihood of an accurate identification when they heard all-inclusive Henderson instructions \( (M = 6.72) \) compared to when they heard case-specific Henderson instructions \( (M = 5.46; F(2, 283) = 2.26, p = .11, \eta^2 = .02) \). Moreover, when evidence quality was good and jurors rendered a verdict prior to evaluating the evidence, there were no significant differences in ratings as a function of instructions \( (all \, ps > .55) \).

**Lineup instructions.** There was a significant main effect of quality of system and estimator variables on ratings of the impact of lineup instructions on the likelihood of an accurate identification \( (F(1, 282) = 141.14, p < .001, \eta^2 = .33) \). Mock jurors rated the lineup instructions as increasing the likelihood of an accurate identification when the quality of the system and
estimator variables was good \( M = 6.61 \) compared to when the quality of the system and estimator variables was poor \( M = 4.05 \). There were no other significant main effects or interactions.

Simple effects tests revealed a significant two-way interaction between evidence quality and evidence evaluation on ratings of the impact of lineup instructions, when evidence evaluation occurred after rendering a verdict \( F (1, 282) = 88.93, p < .001, \eta^2 = .24 \) and when evidence evaluation occurred prior to rendering a verdict \( F (1, 282) = 54.66, p < .001, \eta^2 = .16 \). The larger effect was found when participants evaluated evidence after rendering a verdict, such that they rated the lineup instructions as increasing the likelihood of an accurate identification when evidence quality was good \( M = 6.81 \) compared to when evidence quality was poor \( M = 3.96 \). When participants evaluated evidence prior to rendering a verdict, they rated the lineup instructions as increasing the likelihood of an accurate identification when evidence quality was good \( M = 6.42 \) compared to when evidence quality was poor \( M = 4.13 \). The simple effects tests did not reveal any other significant main effects or interactions.

**Lighting.** There was a significant main effect of quality of system and estimator variables on ratings of the impact of lighting on the likelihood of an accurate identification \( F (1, 283) = 272.62, p < .001, \eta^2 = .49 \). Mock jurors rated the lighting as increasing the likelihood of an accurate identification when the quality of the system and estimator variables was good \( M = 7.37 \) compared to when the quality of the system and estimator variables was poor \( M = 3.19 \). There were no other significant main effects or interactions.

Simple effects tests revealed a significant interaction between evidence quality and evidence evaluation on ratings of the impact of lighting, when evidence evaluation occurred after rendering a verdict \( F (1, 283) = 155.84, p < .001, \eta^2 = .36 \) and when evidence evaluation
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

occurred prior to rendering a verdict \((F(1, 283) = 118.54, p < .001, \eta^2 = .30)\). The larger effect was found when participants evaluated evidence after rendering a verdict, such that they rated the lighting as increasing the likelihood of an accurate identification when evidence quality was good \((M = 7.54)\) compared to when evidence quality was poor \((M = 3.12)\). Similarly, when participants evaluated evidence prior to rendering a verdict, they rated the lighting as increasing the likelihood of an accurate identification when evidence quality was good \((M = 7.20)\) compared to when evidence quality was poor \((M = 3.25)\). The simple effects tests did not reveal any other significant main effects or interactions.

**Exposure duration.** There was a significant main effect of quality of system and estimator variables on ratings of the impact of exposure duration on the likelihood of an accurate identification \((F(1, 283) = 186.44, p < .001, \eta^2 = .40)\). Mock jurors rated the exposure duration as increasing the likelihood of an accurate identification when the quality of the system and estimator variables was good \((M = 6.85)\) compared to when the quality of the system and estimator variables was poor \((M = 3.41)\). There were no other significant main effects or interactions.

Simple effects tests revealed a significant interaction between evidence quality and evidence evaluation on ratings of the impact of exposure duration, when evidence evaluation occurred after rendering a verdict \((F(1, 283) = 107.67, p < .001, \eta^2 = .28)\) and when evidence evaluation occurred prior to rendering a verdict \((F(1, 283) = 80.14, p < .001, \eta^2 = .22)\). Specifically, when participants evaluated evidence after rendering a verdict, they rated the exposure duration as increasing the likelihood of an accurate identification when evidence quality was good \((M = 7.08)\) compared to when evidence quality was poor \((M = 3.44)\). Similarly, when participants evaluated evidence prior to rendering a verdict, they rated the exposure
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

duration as increasing the likelihood of an accurate identification when evidence quality was good ($M = 6.61$) compared to when evidence quality was poor ($M = 3.39$). The simple effects tests did not reveal any other significant main effects or interactions.

**If/Then Statements**

For all analyses, participants answers were coded as such: yes = 1, no/no answer = 0. Thus, for all $\beta$s, higher numbers indicate a change in verdict and lower numbers indicate no change in verdict.

**Lineup instructions.** Participants indicated either ‘yes’, ‘no’, or ‘no answer’ to the following question: “Would your final verdict be different if the police had told the witness that the suspect may or may not be in the lineup?” The overall model was significant ($F(3, 291) = 3.62, p = .01$, adjusted $R^2 = .03$). Both instruction type ($\beta = .13, p = .03$) and evidence quality ($\beta = -.14, p = .02$) significantly predicted whether or not participants reported they would change their verdict if the police provided unbiased lineup instructions to the witness. Specifically, participants who received either the all-inclusive Henderson (39.4%) or the case-specific Henderson (35.2%) instruction were marginally more likely to report that they would change their verdict compared to those who heard no eyewitness instruction (25.4%; $\chi^2 (2) = 4.90, p = .09, \varphi = .13$). Additionally, participants who read about poor eyewitnessing conditions (63.4%) were more likely to report that they would change their verdict compared to those who read about good eyewitnessing conditions (36.6%; $\chi^2 (1) = 5.88, p = .02, \varphi = -.14$).

**Lineup procedure.** Participants indicated either ‘yes’, ‘no’, or ‘no answer’ to the following question: “Would your final verdict be different if the police had not shown the witness both a mugshot lineup and an in-person lineup?” The overall model was significant ($F(3, 291) = 4.44, p < .01$, adjusted $R^2 = .03$). Evidence quality ($\beta = -.17, p < .01$) and evidence
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

evaluation ($\beta = .12$, $p = .04$) significantly predicted whether or not participants reported they would change their verdict if the police had not shown the witness both a mugshot and a lineup. Specifically, participants who read about poor eyewitnessing conditions (66.2%) were more likely to report that they would change their verdict compared to those who read about good eyewitnessing conditions (33.8%; $\chi^2 (1) = 8.31$, $p < .01$, $\varphi = -.17$). Additionally, participants who evaluated the evidence prior to rendering a verdict were more likely (60.3%) to report that they would change their verdict compared to those who evaluated the evidence after rendering a verdict (39.7%; $\chi^2 (1) = 4.39$, $p = .04$, $\varphi = .12$).

**Lighting.** Participants indicated either ‘yes’, ‘no’, or ‘no answer’ to the following question: “Would your final verdict be different if the lighting had been better at the time of the crime?” The overall model was significant ($F(3, 291) = 17.113$, $p < .001$, adjusted $R^2 = .14$). Instruction type ($\beta = .15$, $p < .01$) and evidence quality ($\beta = -.35$, $p < .001$) significantly predicted whether or not participants reported they would change their verdict if the lighting had been better at the time of the crime. Specifically, participants who received either the all-inclusive Henderson (38.5%) or the case-specific Henderson (35.6%) instruction were more likely to report that they would change their verdict compared to those who heard no eyewitness instruction (26.0%; $\chi^2 (2) = 7.07$, $p = .03$, $\varphi = .16$). Additionally, participants who read about poor eyewitnessing conditions (75.0%) were more likely to change their verdict compared to those who read about good eyewitnessing conditions (25.0%; $\chi^2 (1) = 37.49$, $p < .001$, $\varphi = -.36$).

**Duration.** Participants indicated either ‘yes’, ‘no’, or ‘no answer’ to the following question: “Would your final verdict be different if the eyewitness saw the perpetrator for a longer period of time?” The overall model was significant ($F(3, 265) = 8.68$, $p < .001$, adjusted $R^2 = .08$). Instruction type ($\beta = .16$, $p < .01$) and evidence quality ($\beta = -.23$, $p < .001$) significantly
predicted whether or not participants reported they would change their verdict if the eyewitness saw the perpetrator for a longer period of time. Specifically, participants who received the all-inclusive *Henderson* (41.0%) instruction were more likely to report that they would change their verdict compared to those who heard the case-specific *Henderson* (30.3%) or no eyewitness instruction (28.7%; \( \chi^2 (2) = 7.49, p = .02, \phi = .17 \)). Additionally, participants who read about poor eyewitnessing conditions (68.9%) were more likely to report that they would change their verdict compared to those who read about good eyewitnessing conditions (31.1%; \( \chi^2 (1) = 15.51, p < .001, \phi = -.24 \)).

**Discussion**

Results of study 2 confirmed both hypothesis one and hypothesis two. A significant main effect of evidence quality indicated that jurors were able to differentiate between good and poor witnessing conditions without the help of eyewitness jury instructions (\( \beta = .92 \)). The main effect of evidence quality on verdict was qualified by the significance of comparison 1, wherein participants who received no eyewitness instruction and evaluated the evidence after rendering a verdict were more likely to convict when evidence quality was good compared to when evidence quality was poor (\( \beta = 1.13 \)).

Results of the planned comparisons suggest that, while jurors are sensitive to evidence quality without the help of instructions, the sensitivity effects increase with the help of the case-specific *Henderson* instructions and increase even further with the help of the all-inclusive *Henderson* instructions. Specifically, jurors were sensitive to evidence quality when they evaluated the evidence after rendering a verdict and received case-specific *Henderson* instructions (comparison 3: \( \beta = 1.41 \)). The largest sensitivity effect, however, was found when jurors evaluated the evidence prior to rendering a verdict and they received all-inclusive
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

*Henderson* instructions (comparison 6: $\beta = 2.08$). This effect was further supported by the significance of comparison 12, wherein jurors who received all-inclusive *Henderson* instructions were more likely to convict when evidence quality was good compared to poor ($\beta = 1.49$). Thus, evidence evaluation prior to rendering a verdict and non-specific instructions on eyewitness identification increased jurors’ sensitivity to identification quality. It appears that testing may have increased jurors’ memory for the instruction, specifically for the all-inclusive *Henderson* instructions. That is, asking jurors to evaluate the impact of each variable during the judicial instruction, as well as testing them (prior to rendering a verdict) on their memory of the case facts and the impact each manipulated variable had on the identification, appears to be the most effective way to increase sensitivity.

Why were jurors more sensitive to the identification quality when they heard a longer instruction that included more discussion of eyewitness variables that were irrelevant to their specific case? For the case-specific *Henderson* instructions, jurors received instruction on the manipulated variables (procedure, lineup instructions, lighting, and duration) when evidence quality was poor and did not receive instruction on these variables when evidence quality was good. For the all-inclusive *Henderson* instructions, juror received instruction on the manipulated variables *as well as* instruction on all eyewitness variables, regardless of whether evidence quality was good or poor. More specifically, jurors in the poor evidence quality conditions who heard the all-inclusive *Henderson* instruction received judicial instruction on several eyewitness variables that may have decreased the likelihood of an accurate identification (i.e., stress, procedure, lineup instructions, lighting, and duration, weapon focus, time elapsed) as well as several eyewitness variables that may have increased the likelihood of an accurate identification (assuming that if variables were not mentioned in the transcript they were not an issue; distance,
intoxication, disguise/changed appearance, lineup composition, fillers, confirmatory feedback, double-blind). In comparison, jurors in the good evidence quality conditions who received the all-inclusive *Henderson* instructions should have interpreted all of the instructions on eyewitness variables as having *increased* the likelihood of an accurate identification. Thus, we were able to better detect sensitivity to identification quality by including mention of all eyewitness variables in the judicial instruction and in turn drawing jurors’ attention to both the poor and the good eyewitnessing conditions. This is an important finding for courts to consider. Are variables that are not being included in the case-specific instruction actually irrelevant to the case? Are eyewitness instructions more effective when they help jurors by pointing out the poor and the good eyewitnessing conditions?

Participants’ perceptions of some of the evidence did differ as a function of the independent variables. In fact, jurors indicated that each of the evidence perception variables (identification procedure, lineup instructions, lighting, exposure duration) had a greater impact on the likelihood of an accurate identification when the quality of system and estimator variables was good compared to when the quality of system and estimator variables was poor. This effect was in the expected direction, indicating that participants who heard the version of the trial with poor evidence quality reported each evidence perception variable as decreasing the likelihood of an accurate identification.

Participants’ ratings of the impact of the lineup procedure differed as a function of evidence evaluation, such that jurors who evaluated the evidence prior to rendering a verdict rated the lineup procedure as increasing the likelihood of an accurate identification compared to jurors who evaluated the evidence after rendering a verdict. This main effect is relatively unimportant, as it is confounded with the quality of identification evidence. The marginally
significant three-way interaction of the three independent variables on ratings of the impact of lineup procedure indicates a sensible pattern of results. Specifically, when evidence quality was poor and jurors evaluated the evidence prior to rendering a verdict, jurors rated the lineup procedure as decreasing the likelihood of an accurate identification when they heard all-inclusive Henderson instructions, and increasing the likelihood of an accurate identification when they heard no eyewitness instruction. This finding indicates that a combination of all-inclusive Henderson instructions and evaluation of the evidence prior to rendering a verdict may help draw jurors’ attention to the potentially negative impact of the multiple viewings system variable.

Ratings of the impact of the lineup instructions differed as a function of the interaction between evidence quality and evidence evaluation. Specifically, jurors who evaluated the evidence after rendering a verdict rated the lineup instructions as increasing the likelihood of an accurate identification when the evidence quality was good and decreasing the likelihood of an accurate identification when the evidence quality was poor. A similar pattern was found when jurors evaluated the evidence prior to rendering a verdict, such that they rated the lineup instructions as increasing the likelihood of an accurate identification when the evidence quality was good and decreasing the likelihood of an accurate identification when the evidence quality was poor. Although the pattern of results is similar, there was a larger effect for those who evaluated the evidence after rendering a verdict ($\eta^2 = .24$) compared to those who evaluated the evidence prior to rendering a verdict ($\eta^2 = .16$), indicating that evidence evaluation prior to rendering a verdict was not a more helpful alternative to rendering an immediate verdict, at least for this variable.

Ratings of the impact of the lighting at the crime scene differed as a function of the interaction between evidence quality and evidence evaluation. Specifically, jurors who evaluated
the evidence after rendering a verdict rated the lighting as increasing the likelihood of an 
accurate identification when the evidence quality was good and decreasing the likelihood of an 
accurate identification when the evidence quality was poor. A similar pattern was found when 
jurors evaluated the evidence prior to rendering a verdict, such that they rated the lighting as 
increasing the likelihood of an accurate identification when the evidence quality was good and 
decreasing the likelihood of an accurate identification when the evidence quality was poor. 
Similar to the pattern found with the impact of lineup instructions variable, there was a larger 
effect for those who evaluated the evidence after rendering a verdict ($\eta^2 = .36$) compared to those 
who evaluated the evidence prior to rendering a verdict ($\eta^2 = .30$), indicating that evidence 
evaluation prior to rendering a verdict was not a more helpful alternative to rendering an 
immediate verdict, for the lighting variable either.

Ratings of the impact of the duration also differed as a function of the interaction 
between evidence quality and evidence evaluation. Specifically, jurors who evaluated the 
evidence after rendering a verdict rated the duration as increasing the likelihood of an accurate 
identification when the evidence quality was good and decreasing the likelihood of an accurate 
identification when the evidence quality was poor. A similar pattern was found when jurors 
evaluated the evidence prior to rendering a verdict, such that they rated the duration as increasing 
the likelihood of an accurate identification when the evidence quality was good and decreasing 
the likelihood of an accurate identification when the evidence quality was poor. Similar to the 
pattern found with the impact of lineup instructions and lighting variables, there was a larger 
effect for those who evaluated the evidence after rendering a verdict ($\eta^2 = .28$) compared to those 
who evaluated the evidence prior to rendering a verdict ($\eta^2 = .22$), indicating that evidence
evaluation prior to rendering a verdict was not a more helpful alternative to rendering an immediate verdict, for the duration variable either.

Although evidence evaluation prior to rendering a verdict did not seem to affect jurors’ perceptions of the impact of each of the manipulated variables, it does appear that giving jurors the option to decide whether their verdict would have been different if the manipulated variable had been better or worse may help them integrate their knowledge of the instructions into their verdict decisions. Results of the if/then statement analyses show that jurors are knowledgeable enough, with the help of jury instructions, to use counterfactual thinking and sensibly report whether they would render a different verdict if there had been changes in some eyewitnessing conditions. When participants were asked whether they would have changed their verdict had the police given unbiased lineup instructions, participants who heard either of the Henderson instructions were more likely to report ‘yes’ compared to those who heard no eyewitness instruction ($\phi = .13$), although this effect was only marginally significant. Participants who read about poor eyewitnessing conditions were also more likely to report a ‘yes’ compared to those who read about good eyewitnessing conditions ($\phi = -.14$). When participants were asked whether they would have changed their verdict had the police not shown the witness both a mugshot and a lineup, participants who read about poor eyewitnessing conditions were more likely to report a ‘yes’ compared to those who read about good eyewitnessing conditions ($\phi = -.17$). Participants who evaluated the evidence before rendering a verdict were also more likely to change their verdict had the police not shown the witness both a mugshot and a lineup, compared to those who evaluated the evidence after rendering a verdict ($\phi = .12$).

Some of the effects for the estimator variables were larger compared to the system variables. When participants were asked whether their verdict would have been different if the
lighting was better at the scene of the crime, participants who heard either of the *Henderson* instructions were more likely to report ‘yes’ compared to those who heard no eyewitness instruction ($\phi = .16$). Participants who read about poor eyewitnessing conditions were also more likely to report a ‘yes’ compared to those who read about good eyewitnessing conditions ($\phi = - .36$). When participants were asked whether their verdict would have been different if the witness had seen the perpetrator for a longer period of time, participants who heard the all-inclusive *Henderson* instruction were more likely to change their verdict compared to those who heard the case-specific *Henderson* instruction or no eyewitness instruction ($\phi = .17$). Participants who read about poor eyewitnessing conditions were also more likely to report a ‘yes’ compared to those who read about good eyewitnessing conditions ($\phi = -.24$).

All correlations between variables were statistically significant. Correlations were strongest, however, between system variables (lineup procedure & biased instruction: $r = .59$, $p < .001$) and estimator variables (lighting & duration: $r = .71$, $p < .001$). Results of the if/then statement analyses, overall, indicate that jurors are sensitive to the impact of each of the manipulated eyewitnessing conditions as a function of eyewitness jury instructions, evidence evaluation, and evidence quality, and thinking about changes in the manipulated variables (particularly estimator variables) might be enough to move verdicts around. That jurors were not more likely to report that they would change their verdict if certain variables had been different; may indicate that jurors give more weight to the influence of certain variables over others. Future research should explore which variables are more or less influential than others, as well as the possibility that jurors may require a certain number or combination of problematic eyewitnessing variables in order to move from an acquittal to a conviction.
Additionally, future research should tailor the if/then statements to each specific condition. In this study, all participants were given the same if/then scenario (i.e., they were asked their verdict would have been different if the eyewitnessing conditions had been better?), regardless of whether they read about good or poor witnessing conditions. In order to better test whether participants are sensitive to evidence quality, those who see the good witnessing conditions should be asked to imagine a scenario in which the witnessing conditions were poor, and then to render a verdict. Thus, it would be more evident whether participants truly understood the Henderson instructions.

As with any online study, there are certainly limitations with the data collection. It is worth noting, however, that the majority of research indicates that data collected via mTurk is comparable (e.g., Johnson & Borden, 2012) or even superior (e.g., Birnbaum, 2000; Casler, Bickel, & Hackett, 2013) to data collected in person. In an attempt to ameliorate the possibility that participants were not reading the judicial instructions carefully, we included an instruction for participants to type “I have read the judge’s instructions” when prompted on a later page. Thus, we were able to determine that there was no difference in the pattern of results based on participants’ responses to this prompt. Because the study was conducted online, participants had to read the judicial instructions (rather than hear them directly from a judge). This may have affected how seriously the participants took the instructions. Future research should try to replicate this study using a video trial simulation to increase the generalizability of the results. Another limitation with this study is that the system and estimator variables were grouped together, so it was impossible to show whether each variable had an effect independent of the other variables. Future research should attempt to manipulate each variable independently so that
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

researchers can evaluate which system or estimator variables jurors have a more difficult time interpreting.

Overall, study 2 demonstrates that one way to increase jurors’ sensitization to eyewitness identification is to provide jury instructions on all eyewitness variables, regardless of whether or not they are specific to the case. The use of the all-inclusive Henderson instruction (in comparison to the case-specific Henderson instructions) is actually a very practical option for courts, in that it is much less time-consuming and costly to include instruction on all eyewitness variables, regardless of which are relevant to a specific case. That is, inclusion of all eyewitness variables in the judicial instructions does not require an exercise of judicial discretion as to whether certain eyewitnessing conditions are problematic in each case. Moreover, the result that evidence evaluation helps to increase jurors’ sensitivity to identification quality should be included in future jury decision-making research, as evidence evaluation is likely similar to the effects of deliberation.

Overall Discussion

The purpose of the Henderson instructions is to not only inform jurors about relevant system and estimator variables, but to directly instruct jurors about the potential impact of such variables on the accuracy of an eyewitness identification. Results of the first study did not support the Henderson court’s assumption that the instructions will help jurors analyze and evaluate the credibility of eyewitness identification. Whereas there was no evidence of a sensitivity effect, there was evidence that the manipulated variables influenced perceptions of the evidence and witnesses, thus partially supporting our hypothesis that ratings of the witnesses and evidence perceptions would mediate the relationship between the manipulated variables and verdict. There was also evidence of a skepticism effect, such that jurors’ impressions were
influenced by the manipulated variables, but instructions improved neither jurors’ verdict
decisions nor the quality of evidence evaluations.

Results of the second study were more promising, as the results supported both of our
hypotheses. Jurors were sensitive to evidence quality when they evaluated the evidence after
rendering a verdict and received case-specific Henderson instructions. The sensitivity effect was
largest, however, when jurors evaluated the evidence before rendering a verdict and they
received the all-inclusive Henderson instructions. Thus, one key to increasing juror sensitivity
may be to draw jurors’ attention to the good eyewitnessing conditions as well as the poor
eyewitnessing conditions.

Taken together, the current studies bring us closer to understanding which system and
estimator variables jurors need more help to understand. Future research should focus on these
variables. If we are able to determine how to sensitize jurors to each of these variables
independently, we will be closer to developing the most effective eyewitness identification
instructions yet. In particular, future research should look at the correlation between laypersons’
knowledge of eyewitness identification variables (Desmarais & Reade, 2010) and their ability (or
lack thereof) to integrate this knowledge into their decision-making.

Results of the current studies are not entirely dissimilar from previous research, in that
jurors were aware of the effects of some manipulated variables but they often failed to integrate
this knowledge into their verdict decisions. It is important to recognize that the sensitivity effect
expected to occur as a result of these judicial instructions may be quite difficult for jurors (Cutler,
Penrod, & Dexter, 1989). That is, although laypeople may understand some of the eyewitness
issues (e.g., Desmrairs & Read, 2010), if trained judges are not proficient integrators of
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

knowledge (Slovic, 1969; Hoffman, Slovic, & Rorer, 1968), how can the courts expect laypeople to be?

Overall, the results of the current studies suggest that the current Henderson instructions should be modified to improve juror sensitivity to various witnessing and identification conditions. The findings from study 2 represent a promising avenue for the future of the Henderson instructions, as they indicate we are getting closer to discovering the most effective way to instruct jurors about the reliability of eyewitness identification. It is clear from this research that the development of a more concise version of the Henderson instruction is a promising avenue and should be the subject of further research.
References


doi:10.1023/A:1012840831846


doi:10.1037/0021-9010.62.3.311


Experimental Psychology: Applied, 14, 5-20.


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

doi:10.1007/BF01067032


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION


Godfrey, R. D., & Clark, S. E. (2010). Repeated eyewitness identification procedures:
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

doi: 10.1007/s10979-009-9187-7

doi:10.1037/0022-3514.64.4.525


doi:10.1037/a0029779


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Tietjen, T. (2005). *The impact of job related stress on individuals in law enforcement.* Presentation, Fresno Police Department, Sierra Education and Research Institute, and Alliant International University, Fresno, CA.


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION


JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

novel New Jersey jury instruction guiding evaluation of eyewitness testimony, in
progress.
Table 1

*Descriptive Statistics for Verdict (Study 1)*

<table>
<thead>
<tr>
<th>Good or Poor System?</th>
<th>Not Guilty n(%)</th>
<th>Guilty n(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poor</td>
<td>144 (63.3%)</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>129 (55.1%)</td>
</tr>
<tr>
<td>Good or Poor Estimator?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>131 (55.3%)</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>142 (62.3%)</td>
</tr>
<tr>
<td>Instruction Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No EW Instruction</td>
<td>75 (48.4%)</td>
</tr>
<tr>
<td></td>
<td>Henderson-before</td>
<td>101 (67.8%)</td>
</tr>
<tr>
<td></td>
<td>Henderson-after</td>
<td>97 (60.2%)</td>
</tr>
</tbody>
</table>
Table 2

**Effects of Step 1, Step 2, and Step 3 Variables on Verdict (Study 1)**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Variables</td>
<td>.09</td>
<td>.02</td>
<td>.07</td>
</tr>
<tr>
<td>Estimator Variables</td>
<td>-.10*</td>
<td>.05</td>
<td>-.15</td>
</tr>
<tr>
<td>Pre-Henderson</td>
<td>-.18**</td>
<td>-.16**</td>
<td>-.02</td>
</tr>
<tr>
<td>Post-Henderson</td>
<td>-.12*</td>
<td>-.06</td>
<td>-.06</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Henderson X System</td>
<td>-.02</td>
<td>-</td>
<td>-.01</td>
</tr>
<tr>
<td>Pre-Henderson X Estimator</td>
<td>.02</td>
<td>-</td>
<td>.10</td>
</tr>
<tr>
<td>Post-Henderson X System</td>
<td>-.08</td>
<td>-</td>
<td>-.01</td>
</tr>
<tr>
<td>Post-Henderson X Estimator</td>
<td>.06</td>
<td>-</td>
<td>.02</td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact Lineup Instructions</td>
<td>-</td>
<td>-.03</td>
<td>-</td>
</tr>
<tr>
<td>Impact Duration</td>
<td>-</td>
<td>.13**</td>
<td>-</td>
</tr>
<tr>
<td>Impact Weapon</td>
<td>-</td>
<td>.04</td>
<td>-</td>
</tr>
<tr>
<td>Impact Time Delay</td>
<td>-</td>
<td>.06</td>
<td>-</td>
</tr>
<tr>
<td>Impact Identification Procedure</td>
<td>-</td>
<td>.15**</td>
<td>-</td>
</tr>
<tr>
<td>Impact Confirmatory Feedback</td>
<td>-</td>
<td>.05</td>
<td>-</td>
</tr>
<tr>
<td>Detective</td>
<td>-</td>
<td>.11**</td>
<td>-</td>
</tr>
<tr>
<td>Defendant</td>
<td>-</td>
<td>-.27***</td>
<td>-</td>
</tr>
<tr>
<td>Victim</td>
<td>-</td>
<td>.19***</td>
<td>-</td>
</tr>
</tbody>
</table>

*p ≤ .05. **p ≤ .01. ***p ≤ .001.
Table 3

*Effects of Hypothesized Mediators on IVs and Interactions (Study 1)*

<table>
<thead>
<tr>
<th></th>
<th>Sys</th>
<th>Est</th>
<th>Pre</th>
<th>Post</th>
<th>Pre X Sys</th>
<th>Pre X Est</th>
<th>Post X Sys</th>
<th>Post X Est</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Lineup Instructions</td>
<td>.51**</td>
<td>-.10*</td>
<td>-.10*</td>
<td>-.11*</td>
<td>-.13</td>
<td>-.04</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Impact Duration</td>
<td>.07</td>
<td>.15*</td>
<td>-.13*</td>
<td>-.05</td>
<td>-.10</td>
<td>-.12</td>
<td>-.14</td>
<td>.05</td>
</tr>
<tr>
<td>Impact Weapon</td>
<td>.14*</td>
<td>.16**</td>
<td>-.12*</td>
<td>-.07</td>
<td>.03</td>
<td>-.23*</td>
<td>.04</td>
<td>-.11</td>
</tr>
<tr>
<td>Impact Time Delay</td>
<td>.09*</td>
<td>.31**</td>
<td>-.10*</td>
<td>-.10*</td>
<td>-.07</td>
<td>.01</td>
<td>-.05</td>
<td>.06</td>
</tr>
<tr>
<td>Impact ID Procedure</td>
<td>.33**</td>
<td>-.11*</td>
<td>-.17**</td>
<td>-.13*</td>
<td>-.14</td>
<td>.02</td>
<td>-.09</td>
<td>.05</td>
</tr>
<tr>
<td>Impact Feedback</td>
<td>-.56**</td>
<td>.01</td>
<td>.01</td>
<td>-.01</td>
<td>-.07</td>
<td>.10</td>
<td>-.00</td>
<td>.10</td>
</tr>
<tr>
<td>Detective</td>
<td>.17**</td>
<td>-.13*</td>
<td>-.12*</td>
<td>-.10</td>
<td>.05</td>
<td>-.10</td>
<td>-.00</td>
<td>.08</td>
</tr>
<tr>
<td>Defendant</td>
<td>-.11*</td>
<td>.13*</td>
<td>.06</td>
<td>.04</td>
<td>-.04</td>
<td>.09</td>
<td>.09</td>
<td>-.05</td>
</tr>
<tr>
<td>Victim</td>
<td>.13*</td>
<td>-.10*</td>
<td>.00</td>
<td>-.03</td>
<td>.04</td>
<td>-.13</td>
<td>-.06</td>
<td>.03</td>
</tr>
</tbody>
</table>

*p ≤ .05. **p ≤ .001.

*Note.* Sys = system variables; Est = estimator variables; Pre = pre-Henderson instruction dummy variable; Post = post-Henderson instruction dummy variable.
Table 4

Impact of Manipulated Variables on ID Accuracy (Study 1)

<table>
<thead>
<tr>
<th>Instruction Type</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>Stand. Error</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>η² [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No EW Instruction</td>
<td>Poor</td>
<td>Poor</td>
<td>4.23</td>
<td>.28</td>
<td>.18</td>
<td>1,456</td>
<td>.67</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.47</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>[-1.08, .70]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.55</td>
<td>.34</td>
<td>.98</td>
<td>1,456</td>
<td>.32</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>6.11</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[-.43, 1.30]</td>
</tr>
<tr>
<td><em>Henderson</em>-before</td>
<td>Poor</td>
<td>Poor</td>
<td>3.47</td>
<td>.29</td>
<td>.13</td>
<td>1,456</td>
<td>.72</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>3.63</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>[-1.06, .73]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.82</td>
<td>.37</td>
<td>5.67</td>
<td>1,456</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.72</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[.19, 2.02]</td>
</tr>
<tr>
<td><em>Henderson</em>-after</td>
<td>Poor</td>
<td>Poor</td>
<td>3.94</td>
<td>.27</td>
<td>.76</td>
<td>1,456</td>
<td>.38</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>3.55</td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td>[-.49, 1.27]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.41</td>
<td>.33</td>
<td>5.05</td>
<td>1,456</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.44</td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
<td>[-.12, 1.83]</td>
</tr>
</tbody>
</table>
## Duration

| Instruction Type | Good or poor system? | Good or poor estimator? | Means | SE | F  | df | p    | η² [CIs] |
|------------------|----------------------|-------------------------|-------|    |    |    |      |       |
| **No EW Instruction** | Poor | 4.77 | .34 |    | 1.39 | 1,454 | .24 | .003 [-1.69, .42] |
|                   | Good  | 5.40 | .42 |    |     |       |      |       |
|                   | Poor | 4.73 | .40 |    | .66  | 1,454 | .42 | .001 [-1.47, .61] |
|                   | Good  | 5.16 | .34 |    |     |       |      |       |
| **Henderson-before** | Poor | 3.18 | .34 |    | 14.79 | 1,454 | <.001 | .03 [-3.16, -1.02] |
|                   | Good  | 5.27 | .42 |    |     |       |      |       |
|                   | Poor | 4.39 | .44 |    | .96  | 1,454 | .33 | .002 [-1.63, .54] |
|                   | Good  | 4.94 | .34 |    |     |       |      |       |
| **Henderson-after** | Poor | 4.31 | .32 |    | .07  | 1,454 | .79 | <.001 [-1.19, .91] |
|                   | Good  | 4.45 | .43 |    |     |       |      |       |
|                   | Poor | 4.97 | .40 |    | .35  | 1,454 | .56 | .001 [-1.33, .71] |
|                   | Good  | 5.28 | .34 |    |     |       |      |       |
## JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

<table>
<thead>
<tr>
<th>Weapon</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>Stand. Error</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>$\eta^2$ [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No EW Instruction</td>
<td>Poor</td>
<td>Poor</td>
<td>4.77</td>
<td>.30</td>
<td>.18</td>
<td>1,456</td>
<td>.68</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.97</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Poor</td>
<td>5.64</td>
<td>.36</td>
<td>.06</td>
<td>1,456</td>
<td>.81</td>
<td>&lt; .001</td>
<td>[-.80, 1.03]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>5.52</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henderson-before</td>
<td>Poor</td>
<td>Poor</td>
<td>3.64</td>
<td>.31</td>
<td>10.37</td>
<td>1,456</td>
<td>.001</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.20</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Poor</td>
<td>4.39</td>
<td>.39</td>
<td>5.70</td>
<td>1,456</td>
<td>.02</td>
<td>.01</td>
<td>[-2.14, -.21]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>5.57</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henderson-after</td>
<td>Poor</td>
<td>Poor</td>
<td>5.39</td>
<td>.28</td>
<td>1.32</td>
<td>1,456</td>
<td>.25</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.93</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Poor</td>
<td>4.79</td>
<td>.35</td>
<td>2.65</td>
<td>1,456</td>
<td>.10</td>
<td>.006</td>
<td>[-1.65, .16]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>5.54</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

#### Time Delay

<table>
<thead>
<tr>
<th>Instruction Type</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>SE</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>$\eta^2$ [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No EW Instruction</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>3.57</td>
<td>.29</td>
<td>10.56</td>
<td>1,456</td>
<td>.001</td>
<td>.02 [-2.45, -.60]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.10</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>3.82</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>.02 [-2.30, -.50]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.22</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Henderson-before</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>2.93</td>
<td>.30</td>
<td>10.48</td>
<td>1,456</td>
<td>.001</td>
<td>.02 [-2.46, -.60]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.47</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>3.61</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td>.02 [-2.23, -.34]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.89</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Henderson-after</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>3.02</td>
<td>.28</td>
<td>11.32</td>
<td>1,456</td>
<td>.001</td>
<td>.02 [-2.48, -.65]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.59</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>3.91</td>
<td>.35</td>
<td>2.22</td>
<td>1,456</td>
<td>.14</td>
<td>.01 [-1.56, .21]</td>
</tr>
</tbody>
</table>
## Lineup Instructions

<table>
<thead>
<tr>
<th>Instruction Type</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>Stand. Error</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>η² [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No EW Instruction</td>
<td>Poor</td>
<td>Poor</td>
<td>4.23</td>
<td>.28</td>
<td>.18</td>
<td>1,456</td>
<td>.67</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>4.47</td>
<td></td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>[-.08, .70]</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>6.55</td>
<td></td>
<td>.34</td>
<td>.98</td>
<td>1,456</td>
<td>.32</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>6.11</td>
<td></td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[-.43, 1.30]</td>
</tr>
<tr>
<td>Henderson-before</td>
<td>Poor</td>
<td>3.47</td>
<td></td>
<td>.29</td>
<td>.13</td>
<td>1,456</td>
<td>.72</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>3.63</td>
<td></td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>[-1.06, .73]</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>6.82</td>
<td></td>
<td>.37</td>
<td>5.67</td>
<td>1,456</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>5.72</td>
<td></td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[.19, 2.02]</td>
</tr>
<tr>
<td>Henderson-after</td>
<td>Poor</td>
<td>3.94</td>
<td></td>
<td>.27</td>
<td>.76</td>
<td>1,456</td>
<td>.38</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>3.55</td>
<td></td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td>[-.49, 1.27]</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>6.41</td>
<td></td>
<td>.33</td>
<td>5.05</td>
<td>1,456</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>5.44</td>
<td></td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
<td>[-.12, 1.83]</td>
</tr>
</tbody>
</table>
## Identification Procedure

<table>
<thead>
<tr>
<th>Instruction Type</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>SE</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>$\eta^2$ [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No EW Instruction</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>4.94</td>
<td>.30</td>
<td>.02</td>
<td>1,454</td>
<td>.89</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.00</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td>[-.99, .87]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.30</td>
<td>.35</td>
<td>2.59</td>
<td>1,454</td>
<td>.11</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.56</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td>[-.17, 1.66]</td>
</tr>
<tr>
<td><strong>Henderson-before</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>3.69</td>
<td>.30</td>
<td>.25</td>
<td>1,454</td>
<td>.62</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>3.93</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td>[-1.19, -.71]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.14</td>
<td>.38</td>
<td>5.33</td>
<td>1,454</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.02</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td>[.17, 2.08]</td>
</tr>
<tr>
<td><strong>Henderson-after</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>4.29</td>
<td>.28</td>
<td>.58</td>
<td>1,454</td>
<td>.45</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>3.93</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td>[-.57, 1.28]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.03</td>
<td>.35</td>
<td>3.79</td>
<td>1,454</td>
<td>.05</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.15</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[-.01, 1.78]</td>
</tr>
</tbody>
</table>
## Lineup Instructions

<table>
<thead>
<tr>
<th>Instruction Type</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>Stand. Error</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>η² [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>No EW Instruction</td>
<td>Poor</td>
<td>Poor</td>
<td>4.23</td>
<td>.28</td>
<td>.18</td>
<td>1,456</td>
<td>.67</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>4.47</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>[-1.08, .70]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.55</td>
<td>.34</td>
<td>.98</td>
<td>1,456</td>
<td>.32</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>6.11</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[-.43, 1.30]</td>
</tr>
<tr>
<td>Henderson-</td>
<td>Poor</td>
<td>Poor</td>
<td>3.47</td>
<td>.29</td>
<td>.13</td>
<td>1,456</td>
<td>.72</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>before</td>
<td></td>
<td>Good</td>
<td>3.63</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td>[-1.06, .73]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.82</td>
<td>.37</td>
<td>5.67</td>
<td>1,456</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.72</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td>[.19, 2.02]</td>
</tr>
<tr>
<td>Henderson-</td>
<td>Poor</td>
<td>Poor</td>
<td>3.94</td>
<td>.27</td>
<td>.76</td>
<td>1,456</td>
<td>.38</td>
<td>.002</td>
</tr>
<tr>
<td>after</td>
<td></td>
<td>Good</td>
<td>3.55</td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td>[-.49, 1.27]</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.41</td>
<td>.33</td>
<td>5.05</td>
<td>1,456</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.44</td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
<td>[-.12, 1.83]</td>
</tr>
</tbody>
</table>
## Impact of Confirmatory Feedback on Witness Confidence

<table>
<thead>
<tr>
<th>Instruction Type</th>
<th>Good or poor system?</th>
<th>Good or poor estimator?</th>
<th>Means</th>
<th>SE</th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>$\eta^2$ [CIs]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No EW Instruction</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>7.55</td>
<td>.23</td>
<td>1.03</td>
<td>1,451</td>
<td>.31</td>
<td>.002 [-1.12, .36]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>7.93</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>5.30</td>
<td>.28</td>
<td>1.03</td>
<td>1,451</td>
<td>.31</td>
<td>.002 [-1.09, .35]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.67</td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Henderson-before</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>7.54</td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
<td>.001 [-.98, .53]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>7.76</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>6.00</td>
<td>.31</td>
<td>1.57</td>
<td>1,451</td>
<td>.21</td>
<td>.003 [-.28, 1.26]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.51</td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Henderson-after</strong></td>
<td>Poor</td>
<td>Poor</td>
<td>7.64</td>
<td>.22</td>
<td>.11</td>
<td>1,451</td>
<td>.74</td>
<td>&lt;.001 [-.85, .61]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>7.76</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
<td>5.62</td>
<td>.28</td>
<td>.82</td>
<td>1,451</td>
<td>.37</td>
<td>.002 [-.38, 1.03]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>5.29</td>
<td>.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5

*Logistic Regression Overall Model (Study 2)*

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>p</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-specific Henderson</td>
<td>-.50</td>
<td>.29</td>
<td>2.91</td>
<td>1</td>
<td>.09</td>
<td>.61</td>
</tr>
<tr>
<td>All-inclusive Henderson</td>
<td>-.50</td>
<td>.30</td>
<td>2.77</td>
<td>1</td>
<td>.10</td>
<td>.61</td>
</tr>
<tr>
<td>Evidence Quality</td>
<td>.99</td>
<td>.24</td>
<td>16.70</td>
<td>1</td>
<td>&lt;.001</td>
<td>2.70</td>
</tr>
<tr>
<td>Evidence Evaluation</td>
<td>-.12</td>
<td>.24</td>
<td>1.17</td>
<td>1</td>
<td>.28</td>
<td>.88</td>
</tr>
</tbody>
</table>
Table 6

*Verdict as a Function of Evidence Eval., Instruction, & Evidence Quality (Study 2)*

<table>
<thead>
<tr>
<th>Evidence Evaluation</th>
<th>Instruction Type</th>
<th>Evidence Evaluation</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No EW Instruction</td>
<td>Poor</td>
<td>.44</td>
<td>.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>.62</td>
<td>.50</td>
</tr>
<tr>
<td>Pre-Verdict</td>
<td>Case-specific Henderson</td>
<td>Poor</td>
<td>.40</td>
<td>.50</td>
</tr>
<tr>
<td>Verdict Last</td>
<td></td>
<td>Good</td>
<td>.38</td>
<td>.50</td>
</tr>
<tr>
<td></td>
<td>All-inclusive Henderson</td>
<td>Poor</td>
<td>.20</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>.67</td>
<td>.49</td>
</tr>
<tr>
<td>Post-Verdict</td>
<td>No EW Instruction</td>
<td>Poor</td>
<td>.41</td>
<td>.50</td>
</tr>
<tr>
<td>Verdict First</td>
<td></td>
<td>Good</td>
<td>.68</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td>Case-specific Henderson*</td>
<td>Poor</td>
<td>.28</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>.62</td>
<td>.50</td>
</tr>
<tr>
<td></td>
<td>All-inclusive Henderson</td>
<td>Poor</td>
<td>.30</td>
<td>.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td>.54</td>
<td>.51</td>
</tr>
</tbody>
</table>

*p ≤ .05. **p ≤ .01. ***p ≤ .001.
Table 7

_Planned Comparisons of Interest Descriptives_

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison 1</td>
<td>0</td>
<td>1</td>
<td>0 v 1</td>
<td>1.13</td>
<td>3.78</td>
<td>3.09</td>
<td>.10</td>
<td>.05</td>
</tr>
<tr>
<td>Comparison 2</td>
<td>0</td>
<td>0</td>
<td>0 v 1</td>
<td>.71</td>
<td>1.56</td>
<td>2.04</td>
<td>.04</td>
<td>.21</td>
</tr>
<tr>
<td>Comparison 3</td>
<td>1</td>
<td>1</td>
<td>0 v 1</td>
<td>1.41</td>
<td>5.54</td>
<td>4.11</td>
<td>.15</td>
<td>.02</td>
</tr>
<tr>
<td>Comparison 4</td>
<td>1</td>
<td>0</td>
<td>0 v 1</td>
<td>-.07</td>
<td>.01</td>
<td>.94</td>
<td>&lt;.001</td>
<td>.91</td>
</tr>
<tr>
<td>Comparison 5</td>
<td>2</td>
<td>1</td>
<td>0 v 1</td>
<td>.99</td>
<td>2.65</td>
<td>2.70</td>
<td>.08</td>
<td>.10</td>
</tr>
<tr>
<td>Comparison 6</td>
<td>2</td>
<td>0</td>
<td>0 v 1</td>
<td>2.08</td>
<td>8.65</td>
<td>8.00</td>
<td>.28</td>
<td>.003</td>
</tr>
<tr>
<td>Comparison 7</td>
<td>1 v 0</td>
<td>0</td>
<td>0</td>
<td>-.16</td>
<td>.08</td>
<td>.85</td>
<td>.002</td>
<td>.78</td>
</tr>
<tr>
<td>Comparison 8</td>
<td>2 v 0</td>
<td>0</td>
<td>0</td>
<td>-1.15</td>
<td>3.18</td>
<td>.32</td>
<td>.09</td>
<td>.08</td>
</tr>
<tr>
<td>Comparison 9</td>
<td>2 v 1</td>
<td>0</td>
<td>0</td>
<td>.98</td>
<td>2.31</td>
<td>2.67</td>
<td>.07</td>
<td>.13</td>
</tr>
<tr>
<td>Comparison 10</td>
<td>1</td>
<td>n/a</td>
<td>0 v 1</td>
<td>.67</td>
<td>2.66</td>
<td>1.95</td>
<td>.04</td>
<td>.10</td>
</tr>
<tr>
<td>Comparison 11</td>
<td>1 v 0</td>
<td>n/a</td>
<td>0</td>
<td>-.36</td>
<td>.77</td>
<td>.70</td>
<td>.02</td>
<td>.38</td>
</tr>
<tr>
<td>Comparison 12</td>
<td>2</td>
<td>n/a</td>
<td>0 v 1</td>
<td>1.49</td>
<td>10.41</td>
<td>4.42</td>
<td>.16</td>
<td>.001</td>
</tr>
<tr>
<td>Comparison 13</td>
<td>2 v 0</td>
<td>n/a</td>
<td>0</td>
<td>-.78</td>
<td>3.23</td>
<td>.46</td>
<td>.05</td>
<td>.07</td>
</tr>
<tr>
<td>Comparison 14</td>
<td>1 v 2</td>
<td>n/a</td>
<td>0</td>
<td>-.36</td>
<td>.77</td>
<td>.70</td>
<td>.02</td>
<td>.38</td>
</tr>
</tbody>
</table>

**Note.** Inst. = Instruction type (0 = No EW instruction, 1 = Case-specific Henderson, 2 = All-inclusive Henderson); Evid. Eval. = Evidence Evaluation (0 = before verdict, 1 = after verdict); Evid. Qual. = Evidence Quality (0 = poor, 1 = good).
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Figure 1

Study 2 Instruction Manipulation

- **No Instruction**
  - Good EW conditions → Burden of proof, PBRD
  - Poor EW conditions → Burden of proof, PBRD

- **Case-specific Henderson**
  - Good EW conditions → Stress, c/a, double-blind, pdp
  - Poor EW conditions → Stress, c/a, double-blind, pdp, duration, lighting, multiple viewings, lineup instructions

- **All-inclusive Henderson**
  - Good EW conditions → Stress, c/a, double-blind, pdp, duration, lighting, mult. viewings, instructions, weapon, time elapsed, distance, intoxication, disguise, lineup comp., fillers, feedback
  - Poor EW conditions → Stress, c/a, double-blind, pdp, duration, lighting, mult. viewings, instructions, weapon, time elapsed, distance, intoxication, disguise, lineup comp., fillers, feedback

*Note. PBRD = proof beyond a reasonable doubt; c/a = confidence/accuracy; pdp = prior description of the perpetrator; lineup comp. = lineup composition*
Appendix A

*Study 1 No Eyewitness Instruction*

**Instructions to the Jury**

The Court instructs you as follows: under the evidence presented to you in this case, you may find Mr. William Johnson guilty or not guilty of the following offense: criminal attempt rape in the first degree of Marie Sands.

**Criminal attempt rape in the first degree of Marie Sands**

You will find Mr. Johnson guilty of this offense if and only if you believe from the evidence beyond a reasonable doubt all of the following:

A) That in Westchester county, on July 18th, 2011, Mr. Johnson used forcible compulsion against Marie Sands, and
B) That in using forcible compulsion it was his intention to engage in sexual intercourse with Ms. Sands, and
C) That Mr. Johnson's actions constituted a substantial step in the course of conduct planned to result in having sexual intercourse with Ms. Sands.

If you find Mr. Johnson guilty under this instruction, you shall say so by verdict and no more.

This defendant, as are all defendants in criminal cases, is presumed to be innocent until proven guilty beyond a reasonable doubt. That presumption continues throughout the whole trial of the case and even during your deliberations unless and until you have determined that the State has proven his guilt beyond a reasonable doubt. The prosecution must prove its case by more than a mere preponderance of the evidence, yet not necessarily to an absolute certainty. The State has the burden of proving the defendant guilty beyond a reasonable doubt.

Some of you may have served as jurors in civil cases, where you were told that it is necessary to prove only that a fact is more likely true than not true. In criminal cases, the State’s proof must be more powerful than that. It must be beyond a reasonable doubt.

A reasonable doubt is an honest and reasonable uncertainty in your minds about the guilt of the defendant after you have given full and impartial consideration to all of the evidence. A reasonable doubt may arise from the evidence itself or from a lack of evidence. It is a doubt that a reasonable person hearing the same evidence would have. Proof beyond a reasonable doubt is proof, for example, that leaves you firmly convinced of the defendant's guilt. In this world, we know very few things with absolute certainty. In criminal cases the law does not require proof that overcomes every possible doubt.

If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If, on the other hand, you are not firmly convinced of defendant's guilt, you must give defendant the benefit of the doubt and find him not guilty.
Appendix B

*Study 1 Henderson Instruction (most extensive version)*

**Instructions to the Jury**

The Court instructs you as follows:

Mr. William Johnson 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 Marie Sands. You will recall that this witness identified the defendant in court as the person who committed first degree burglary and first degree criminal attempt rape. The State also presented testimony that on a prior occasion before this trial, this witness identified the defendant as the person who committed this offense. According to the witness, her identification of the defendant was based upon the observations and perceptions that she made of the perpetrator at the time the offense was 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 offenses charged.

Eyewitness identifications 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 a mistake in identifications. That research has focused on the nature of memory and the factors that affect the reliability of witness’s 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
**JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION**

- **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 determining whether the eyewitness identification evidence is reliable.

In evaluating misidentification, you should consider the observations and perceptions on which the identification was based, the witness’s ability to make those observations and perceive those 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 when 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 indication of 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 the defendant is the person who committed the offenses 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 related to the witness, the alleged perpetrator, and the criminal incident itself. In particular you should consider 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 the person who committed the offense at the time of the offense and the witness’s degree of attention to the perpetrator of the time of the events.

In making this assessment you should consider the following:

- **Stress** – Even under the best viewing conditions, high levels of stress can reduce the 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.

- **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, or 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.

- **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 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 the presence of the weapon and focus on other details.

- **Lighting** – Inadequate light can reduce the reliability of an identification. You should consider the lighting conditions present at the time of the alleged crime in this case.

- **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 their prior description of the perpetrator.

- **Confidence and accuracy** – You heard testimony that Marie Sands made a statement at the time she identified the defendant asserting her level of certainty that the person she selected is in fact the person that committed the crime. As I explained earlier, a witness’s level of confidence standing alone may not be an identification of the reliability of the identification. Although some research has found that highly confident witnesses are more likely to make an accurate identification, eyewitness confidence is generally an unreliable indicator of accuracy.

- **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.

- **Cross-racial effects** – Research has shown that people may have greater difficulty 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.

In evaluating the reliability of the 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:

- **Showups** – In this case, the witness identified the defendant during a showup, that is, the defendant was the only person showed 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 2 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 they should not be compelled to make an identification. In determining whether the identification is reliable or the result of unduly suggestive procedure you should consider how much time elapsed since witness saw the perpetrator, whether appropriate instructions were given and all circumstances surrounding the showup.

- 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. 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 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 the witness should not feel compelled to make an identification. 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.
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.

You may consider whether the witness was exposed to opinions, descriptions, or identifications given by other witnesses, to photographs, newspaper accounts, or to any other information or influence that may have affected the independence of 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 of 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 was made by a witness as 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, a product of an impression gained at the in court 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 the evidence you determine that the state has not proven beyond a reasonable doubt that Mr. William Johnson was the person who committed these offenses then you must find him not guilty. If, on the other hand, after consideration of all the evidence you are convinced beyond a reasonable doubt that Mr. William Johnson was correctly identified you will then consider whether the State has proven each and every element of the offense’s charge beyond a reasonable doubt.
Under the evidence presented to you in this case, you may find Mr. William Johnson guilty or not guilty of the following offense: criminal attempt rape in the first degree of Marie Sands.

**Criminal attempt rape in the first degree of Marie Sands**

You will find Mr. Johnson guilty of this offense if and only if you believe from the evidence beyond a reasonable doubt all of the following:

D) That in Westchester county, on July 18th, 2011, Mr. Johnson used forcible compulsion against Marie Sands, and

E) That in using forcible compulsion it was his intention to engage in sexual intercourse with Ms. Sands, and

F) That Mr. Johnson's actions constituted a substantial step in the course of conduct planned to result in having sexual intercourse with Ms. Sands.

If you find Mr. Johnson guilty under this instruction, you shall say so by verdict and no more.

This defendant, as are all defendants in criminal cases, is presumed to be innocent until proven guilty beyond a reasonable doubt. That presumption continues throughout the whole trial of the case and even during your deliberations unless and until you have determined that the State has proven his guilt beyond a reasonable doubt. The prosecution must prove its case by more than a mere preponderance of the evidence, yet not necessarily to an absolute certainty. The State has the burden of proving the defendant guilty beyond a reasonable doubt.

Some of you may have served as jurors in civil cases, where you were told that it is necessary to prove only that a fact is more likely true than not true. In criminal cases, the State’s proof must be more powerful than that. It must be beyond a reasonable doubt.

A reasonable doubt is an honest and reasonable uncertainty in your minds about the guilt of the defendant after you have given full and impartial consideration to all of the evidence. A reasonable doubt may arise from the evidence itself or from a lack of evidence. It is a doubt that a reasonable person hearing the same evidence would have. Proof beyond a reasonable doubt is proof, for example, that leaves you firmly convinced of the defendant's guilt. In this world, we know very few things with absolute certainty. In criminal cases the law does not require proof that overcomes every possible doubt.

If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If, on the other hand, you are not firmly convinced of defendant's guilt, you must give defendant the benefit of the doubt and find him not guilty.
Appendix C

Study 1 Post-Trial Questionnaire

POST-TRIAL QUESTIONNAIRE (A)

1. Do you find the defendant, Mr. William Johnson guilty or not guilty of first degree attempted rape? (CIRCLE ONE)

<table>
<thead>
<tr>
<th>Guilty</th>
<th>Not Guilty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. How confident are you in your verdict?

<table>
<thead>
<tr>
<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>Not at all</td>
<td>Confident</td>
<td>Extremely Confident</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Give a percentage between 0% and 100%. What is the probability that the defendant, Mr. William Johnson, committed first degree attempted rape? _____________%

4. Give a percentage between 0% and 100%. What is the probability that the victim, Ms. Marie Sands, correctly identified Mr. William Johnson as the perpetrator? _____________%

5. How confident was the victim, Ms. Marie Sands, in her identification of the defendant?

<table>
<thead>
<tr>
<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>Not at all</td>
<td>Confident</td>
<td>Extremely Confident</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please indicate whether you agree or disagree with the following statements by choosing the number on the scale corresponding to your belief.

1. The defense had a very strong case.

<table>
<thead>
<tr>
<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>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. The prosecution had a very strong case.

<table>
<thead>
<tr>
<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>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Next, we would like you to describe your impressions of witnesses that you heard using a series of adjective pairs. The scales are designed so that you can express the degree to which the person that you are rating seems to fit one end of the scale or the other. Which space you check should depend on the degree to which the word describes the person you are rating.

For example, if you thought that Jane was slightly tall, you would mark the item as follows:

Tall : X : Short

However, if you thought that Jane was extremely short, you should place the “X” next to short:

Tall : X : Short

If you have any questions about this task, you may ask them at this time.

Based on the testimony that you heard, carefully rate your impressions of Detective John Roberts as best you can on each of the following dimensions:

trustworthy : untrustworthy
honest : dishonest
not believable : believable
convincing : unconvincing
certain : uncertain

Based on the testimony that you heard, carefully rate your impressions of Ms. Marie Sands (the victim) as best you can on each of the following dimensions:

trustworthy : untrustworthy
honest : dishonest
not believable : believable
convincing : unconvincing
certain : uncertain
Based on the testimony that you heard, carefully rate your impressions of Mr. Tom Katz (the forensic serologist) as best you can on each of the following dimensions:

- Trustworthy: ____________________________: Untrustworthy
- Honest: _______________________________: Dishonest
- Not believable: ________________________: Believable
- Convincing: ___________________________: Unconvincing
- Certain: _______________________________: Uncertain

Based on the testimony that you heard, carefully rate your impressions of Mr. William Johnson (the defendant) as best you can on each of the following dimensions:

- Trustworthy: ____________________________: Untrustworthy
- Honest: _______________________________: Dishonest
- Not believable: ________________________: Believable
- Convincing: ___________________________: Unconvincing
- Certain: _______________________________: Uncertain

Based on the testimony that you heard, carefully rate your impressions of Ms. Diana LaFonte (the defendant's friend) as best you can on each of the following dimensions:

- Trustworthy: ____________________________: Untrustworthy
- Honest: _______________________________: Dishonest
- Not believable: ________________________: Believable
- Convincing: ___________________________: Unconvincing
- Certain: _______________________________: Uncertain
Please indicate how each witness' testimony influenced your perception of the evidence.

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

1. To what extent did the testimony of **Detective John Roberts** point to the defendant’s innocence or guilt?

   - Strong Evidence
   - For Innocence
   - For Guilt

2. To what extent did the testimony of **Ms. Marie Sands (the victim)** point to the defendant’s innocence or guilt?

   - Strong Evidence
   - For Innocence
   - For Guilt

3. To what extent did the testimony of **Mr. William Johnson (the defendant)** point to the defendant’s innocence or guilt?

   - Strong Evidence
   - For Innocence
   - For Guilt

4. To what extent did the testimony of **Ms. Diana LaFonte (the defendant's friend)** point to the defendant’s innocence or guilt?

   - Strong Evidence
   - For Innocence
   - For Guilt
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Please indicate your responses to the following questions.

1. The judge’s instructions helped me understand the law.
   1 2 3 4 5 6 7 8 9
   Strongly Disagree Strongly Agree

2. The judge’s instructions helped me to evaluate the eyewitness testimony.
   1 2 3 4 5 6 7 8 9
   Strongly Disagree Strongly Agree

3. Eyewitness identification is generally accurate.
   1 2 3 4 5 6 7 8 9
   Strongly Disagree Strongly Agree

POST-TRIAL QUESTIONNAIRE (B)

Please answer the following questions.

1. How much stress did the victim experience during the crime?
   O Extreme stress
   O Moderate stress
   O No stress

2. What impact did the stress experienced by the victim have on the likelihood of an accurate identification?
   1 2 3 4 5 6 7 8 9
   Strongly reduced Accuracy Strongly increased Accuracy

3. What impact did the victim being Caucasian and the perpetrator being African American have on the likelihood of an accurate identification?
   1 2 3 4 5 6 7 8 9
   Strongly reduced Accuracy Strongly increased Accuracy

118
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

4. Did the police inform the victim that the suspect may or may not be present in the identification procedure?
   ○ Yes
   ○ No

5. What impact did the lineup instructions provided by the police to the victim have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<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>Strongly reduced Accuracy</td>
<td>Strongly increased Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. About how much time did the victim indicate she had to see the perpetrator’s face?
   ○ 10-15 seconds
   ○ Around 45 seconds
   ○ 1-2 minutes
   ○ 3-5 minutes

7. What impact did the amount of time the victim had to view the perpetrator’s face have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<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>Strongly reduced Accuracy</td>
<td>Strongly increased Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. What weapon, if any, did the perpetrator use?
   ○ Gun
   ○ None
   ○ Knife
   ○ Hammer

9. What impact did the presence or absence of a weapon have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<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>Strongly reduced Accuracy</td>
<td>Strongly increased Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

10. What was the time delay between when the crime occurred and when the victim identified the defendant?

○ One day
○ Five days
○ One week
○ One month

11. What impact did the time delay between when the crime occurred and when the victim identified the defendant have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<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>Strongly reduced Accuracy</td>
<td>Strongly increased Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12. What were the lighting conditions at the time of the crime?

○ Unknown
○ Bright, good lighting conditions
○ Dark, bad lighting conditions
○ Some lighting, not completely dark

13. What impact did the lighting during the commission of the crime have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<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>Strongly reduced Accuracy</td>
<td>Strongly increased Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14. What was the distance between the victim and the perpetrator during the commission of the crime?

○ 11-15 feet
○ 6-10 feet
○ 1-5 feet
○ Less than 1 foot away
15. What impact did the distance between the perpetrator and the victim have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly increased Accuracy</td>
<td></td>
</tr>
</tbody>
</table>

16. What type of police procedure was used to assist the victim in making an identification?

- Photo lineup
- Showup
- Sketch composite

17. What impact did the identification procedure used to secure the identification have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly increased Accuracy</td>
<td></td>
</tr>
</tbody>
</table>

18. What, if anything, did the police tell the victim after she identified the defendant?

- Nothing.
- Please try again.
- Good, you identified the suspect.
- Would you like something to drink?

19. Do you think statements made by the police to the victim after her identification affected her confidence that she made an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced Confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly increased Confidence</td>
<td></td>
</tr>
</tbody>
</table>
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Please answer the following question about the judge's instructions:

1. Which of the following topics did the judge's instructions cover (check as many as are applicable):
   - Burden of proof (The State has the burden of proving the defendant guilty)
   - Lineup instructions (Telling the witness the suspect may or may not be in the lineup)
   - Duration (Amount of time an eyewitness has to observe an event)
   - Distance (between the eyewitness and perpetrator)
   - Reasonable doubt (Honest and reasonable uncertainty)
   - Confirmatory feedback (Telling the witness they identified the suspect)
   - Weapon focus (Presence of weapon)
   - Time elapsed (Delays between commission of a crime and time identification is made)
   - Showup (The defendant was the only person shown to the witness during identification)
   - Lighting (Lighting conditions present at the time of event)
   - Cross race (the witness identified a suspect of a different race)

Please answer the following questions.

1. Criminals should be caught and convicted by “any means necessary.”

   1  2  3  4  5
   Strongly Disagree  Agree

2. Police routinely lie to protect other police officers.

   1  2  3  4  5
   Strongly Disagree  Agree

3. Defense lawyers don’t really care about guilt or innocence; they are just in business to make money.

   1  2  3  4  5
   Strongly Disagree  Agree

122
4. At trial, an eyewitness’ confidence is a good predictor of his or her accuracy in identifying the defendant as the perpetrator of the crime.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. A police officer who knows which member of the lineup or photo array is the suspect should not conduct the lineup or photo array.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Eyewitness testimony about an event often reflects not only what a witness actually saw, but information obtained later on from other witnesses, the police, the media etc.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. Only in exceptional circumstances should a defendant be convicted of a crime solely on the basis of eyewitness testimony.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix D

Study 2 No Eyewitness Instruction

This defendant, as are all defendants in criminal cases, is presumed to be innocent until proven guilty beyond a reasonable doubt. The prosecution must prove its case by more than a mere preponderance of the evidence, yet not necessarily to an absolute certainty. The State has the burden of proving the defendant guilty beyond a reasonable doubt. A reasonable doubt is an honest and reasonable uncertainty in your minds about the guilt of the defendant after you have given full and impartial consideration to all of the evidence. A reasonable doubt may arise from the evidence itself or from a lack of evidence. If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If, on the other hand, you are not firmly convinced of defendant's guilt, you must give defendant the benefit of the doubt and find him not guilty.
Appendix E

Study 2 Case-Specific Henderson Instruction (most extensive version)

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? Did this enhance or impair 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? Did this enhance or impair her accuracy?

(c) 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? Did this enhance or impair 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. Did the prior description match the photo or person picked out later? Did the prior description provide details or was it just general in nature? Was the witness's testimony at trial 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) 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?
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

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) **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.

Ask yourself: Could the eyewitness have remembered the defendant from a prior viewing (e.g. mugshot) rather than from the scene of the crime? Does this enhance or impair her accuracy?

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: Might the officer conducting the lineup do anything to suggest to the eyewitness which person or picture to choose, or not to choose? Did this strengthen 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? Did this strengthen or weaken the evidence against the defendant?

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.

This defendant, as are all defendants in criminal cases, is presumed to be innocent until proven guilty beyond a reasonable doubt. The prosecution must prove its case by more than a mere preponderance of the evidence, yet not necessarily to an absolute certainty. The State has the burden of proving the defendant guilty beyond a reasonable doubt. A reasonable doubt is an honest and reasonable uncertainty in your minds about the guilt of the defendant after you have given full and impartial consideration to all of the evidence. A reasonable doubt may arise from the evidence itself or from a lack of evidence. If, based on your consideration of the evidence,
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If, on the other hand, you are not firmly convinced of defendant's guilt, you must give defendant the benefit of the doubt and find him not guilty.
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Appendix F

Study 2 All-Inclusive Henderson Instruction

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:

(d) 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? Did this enhance or impair her accuracy?

(e) 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? Did this enhance or impair her accuracy?

(f) 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. 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? Did this enhance or impair her accuracy?

(g) 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? Did this enhance or impair her accuracy?

(h) 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.
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

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

(i) **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? Did this enhance or impair her accuracy?

(j) **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? Did this enhance or impair 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. Did the prior description match the photo or person picked out later? Did the prior description provide details or was it just general in nature? Was the witness's testimony at trial 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? Did this enhance or impair 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? Did this strengthen 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? Did this strengthen 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.

(c) **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: Might the officer conducting the lineup do anything to suggest to the eyewitness which person or picture to choose, or not to choose? Did this strengthen or weaken the evidence against the defendant?

(d) **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? Did this strengthen or weaken the evidence against the defendant?
(3) **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?**

**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 strengthen 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.

This defendant, as are all defendants in criminal cases, is presumed to be innocent until proven guilty beyond a reasonable doubt. The prosecution must prove its case by more than a mere preponderance of the evidence, yet not necessarily to an absolute certainty. The State has
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

the burden of proving the defendant guilty beyond a reasonable doubt. A reasonable doubt is an honest and reasonable uncertainty in your minds about the guilt of the defendant after you have given full and impartial consideration to all of the evidence. A reasonable doubt may arise from the evidence itself or from a lack of evidence. If, based on your consideration of the evidence, you are firmly convinced that the defendant is guilty of the crime charged, you must find him guilty. If, on the other hand, you are not firmly convinced of defendant's guilt, you must give defendant the benefit of the doubt and find him not guilty.
Appendix G

Study 2 Post-Trial Questionnaire (verdict moved to end for Pre-verdict conditions)

POST-TRIAL QUESTIONNAIRE

1. Do you find the defendant, Mr. Christopher Bryan guilty or not guilty of armed robbery? (CIRCLE ONE)
   Guilty Not Guilty

2. How confident are you in your verdict?
   1 2 3 4 5 6 7 8 9
   Not at all Extremely
   Confident

3. Give a percentage between 0% and 100%. What is the probability that the defendant, Mr. Christopher Bryan, committed armed robbery? ____________%

4. Give a percentage between 0% and 100%. What is the probability that the victim, Ms. Jennifer Stone, correctly identified Mr. Christopher Bryan as the perpetrator? ____________%

5. Give a percentage between 0% and 100%. What probability of a correct identification would YOU require in order to convict? ____________%

6. How confident was the victim, Ms. Jennifer Stone, in her identification of the defendant?
   1 2 3 4 5 6 7 8 9
   Not at all Extremely
   Confident

Please indicate whether you agree or disagree with the following statements by choosing the number on the scale corresponding to your belief.

1. The defense had a very strong case.
   1 2 3 4 5 6 7 8 9
   Strongly Disagree
   Strongly Agree
2. The prosecution had a very strong case.

1 2 3 4 5 6 7 8 9
Strongly Disagree

Please indicate how each witness’ testimony influenced your perception of the evidence.

1. To what extent did the testimony of Ms. Jennifer Stone (the victim) point to the defendant’s innocence or guilt?

1 2 3 4 5 6 7 8 9
Strong Evidence For Innocence

2. To what extent did the testimony of Detective Michael Thomas point to the defendant’s innocence or guilt?

1 2 3 4 5 6 7 8 9
Strong Evidence For Innocence

3. To what extent did the testimony of Ms. Christina Hall (the defendant’s girlfriend) point to the defendant’s innocence or guilt?

1 2 3 4 5 6 7 8 9
Strong Evidence For Innocence

Please indicate your responses to the following questions.

3. The judge’s instructions helped me understand the law.

1 2 3 4 5 6 7 8 9
Strongly Disagree

4. The judge’s instructions helped me to evaluate the eyewitness testimony.

1 2 3 4 5 6 7 8 9
Strongly Disagree
3. Eyewitness identification is generally accurate.

   1   2   3   4   5   6   7   8   9
Strongly Disagree Strongly Agree

**Please answer the following questions.**

1. How much stress did the victim experience during the crime?
   - Extreme stress
   - Moderate stress
   - No stress
   - No answer

2. What impact did the stress experienced by the victim have on the likelihood of an accurate identification?

   1   2   3   4   5   6   7   8   9
Strongly reduced Strongly Increased
Accuracy Accuracy

3. Did the police inform the victim that the suspect may or may not be present in the identification procedure?
   - No
   - Yes
   - No answer

4. What impact did the lineup instructions provided by the police to the victim have on the likelihood of an accurate identification?

   1   2   3   4   5   6   7   8   9
Strongly reduced Strongly Increased
Accuracy Accuracy

5. About how much time did the victim indicate she had to see the perpetrator’s face?
   - 10-15 seconds
   - Around 45 seconds
   - 1-2 minutes
   - 3-5 minutes
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

6. What impact did the amount of time the victim had to view the perpetrator’s face have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly Increased Accuracy</td>
</tr>
</tbody>
</table>

7. What was the time delay between when the crime occurred and when the victim identified the defendant?

- One day
- Five days
- One week
- One month

8. What impact did the time delay between when the crime occurred and when the victim identified the defendant have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly Increased Accuracy</td>
</tr>
</tbody>
</table>

9. What were the lighting conditions at the time of the crime?

- Unknown
- Bright, good lighting conditions
- Dark, bad lighting conditions

10. What impact did the lighting during the commission of the crime have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly Increased Accuracy</td>
</tr>
</tbody>
</table>

11. What was the distance between the victim and the perpetrator during the commission of the crime?

- 11-15 feet
- 6-10 feet
- 1-5 feet
- Less than 1 foot away
12. What impact did the distance between the perpetrator and the victim have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly Increased</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td></td>
</tr>
</tbody>
</table>

13. What type of police procedure was used to assist the victim in making an identification?

- Mugshot Photo Lineup only
- In-Person Lineup only
- Both Mugshot Photo Lineup and In-Person Lineup

14. What impact did the identification procedure used to secure the identification have on the likelihood of an accurate identification?

<table>
<thead>
<tr>
<th></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></td>
<td>Strongly reduced</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly Increased</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td></td>
</tr>
</tbody>
</table>

Please answer the following question about the judge's instructions:

15. Which of the following topics did the judge's instructions cover (check as many as are applicable):

- Burden of proof (The State has the burden of proving the defendant guilty)
- Lineup instructions (Telling the witness the suspect may or may not be in the lineup)
- Duration (Amount of time an eyewitness has to observe an event)
- Distance (between the eyewitness and perpetrator)
- Reasonable doubt (Honest and reasonable uncertainty)
- Confirmatory feedback (Telling the witness they identified the suspect)
- Weapon focus (Presence of weapon)
- Time elapsed (Delays between commission of a crime and time identification is made)
- Showup (The defendant was the only person shown to the witness during identification)
- Lighting (Lighting conditions present at the time of event)
- Cross race (the witness identified a suspect of a different race)
JUDICIAL INSTRUCTION ON EYEWITNESS IDENTIFICATION

Please answer the following questions.

16. Would your final verdict be different if the police had told the witness that the suspect may or may not be in the lineup?
   ○ Yes
   ○ No

17. Would your final verdict be different if the police had not shown the witness both a mugshot lineup and an in-person lineup?
   ○ Yes
   ○ No

18. Would your final verdict be different if the lighting had been better at the time of the crime?
   ○ Yes
   ○ No

19. Would your final verdict be different if the eyewitness saw the perpetrator for a longer period of time?
   ○ Yes
   ○ No