2014

Juror Perceptions of Eyewitness Identification Evidence

Timothy G. Wykes
Wilfrid Laurier University, wyke5360@mylaurier.ca

JUROR PERCEPTIONS OF EYEWITNESS IDENTIFICATION EVIDENCE

by

Timothy G. Wykes

B.A. (Honours), Criminology & Justice, University of Ontario Institute of Technology, 2010

THESIS

Submitted to the Faculty of Human & Social Sciences

in partial fulfilment of the requirements for

Master of Arts

Wilfrid Laurier University

Abstract

Jurors rely on eyewitness testimony in deciding a defendant's guilt or innocence. Archival analyses of hundreds of post-conviction DNA exonerations have identified eyewitness misidentification as the highest individual factor contributing to wrongful convictions (Innocence Project, 2014). Internationally, criminal justice systems have employed procedural safeguards (PSs) to educate juries on factors affecting eyewitness identification accuracy. Two such safeguards include the introduction of eyewitness expert testimony during trial proceedings and the reading of cautionary instructions by a presiding judge. In an independent factorial design, this research sought to examine the effects of a model judicial caution drafted by the Ontario Judicial Council (2012) and eyewitness expert testimony on jurors. Viewing court transcripts surrounding a fictional robbery case, jurors were presented with independently varied evidence of eyewitness testimony (low confidence vs. high confidence) and photoarray lineups (no lineup vs. unbiased lineup vs. biased lineup). Numerous juror perceptions were measured, including verdicts, eyewitness credibility, defence case strength, and understanding of the trial. Results indicated that jurors were not unduly influenced by eyewitness confidence or expert testimony. Jurors were more likely to convict the defendant when the eyewitness made an identification from an unbiased lineup compared to a biased lineup. While jurors were able to detect foil bias, evidence of juror confusion was found with respect to lineup fairness ratings from jurors exposed to the judicial caution. Results are summarized and discussed in view of current Canadian trial proceedings.

Acknowledgements

To Dr. Jennifer Lavoie: Thank you for your guidance, determination, and steadfast commitment. You have selflessly imparted your wisdom to me and I am privileged and forever grateful to have written under your mentorship. Continue to inspire scholars in the future. Ancora imparo.

To Dr. Judy Eaton, Dr. Dan Antonowicz, and Dr. Chris Alksnis: I am honoured to have had each of you as members of my thesis examination committee. Thank you for your genuine interest, enthusiasm, and constructive assessments towards this project.

To the Wilfrid Laurier University Department of Criminology: Remnants of shared conversations, lectures, seminars, conferences, and assignments are bestrewn throughout this thesis, my academic endeavors, and my personal character. Thank you for providing me with a fulfilling academic experience. Thank you to department members Dr. Thomas Fleming, Dr. Stacy Hannem, and Marg Harris, for providing advice and unwavering support.

To Dr. Brian Cutler, Mr. Dan Walters, and Mrs. Gail Haynes: Your early encouragement to pursue graduate studies will never be forgotten and always appreciated.

To my family and loved ones: Thank you for your love and patience.

Table of Contents

List of Tables

List of Figures

Introduction

Scholars have long suspected that accounts provided by witnesses to criminal events can be unreliable. On the intersection of the practices of law and psychology, Hugo Münsterberg (1908) argued that "justice would less often miscarry if all who are to weigh evidence were more conscious of the treachery of human memory" (para. 5). However, this polemic was echoed only fleetingly throughout the early- to mid-20th century (e.g., Smith, 1930; Moore, 1949). Research into the fallibility of memory conducted in the 1970's (e.g. Buckhout, 1976; Loftus, 1974; Wells, 1978; Yarmey, 1979) reached unforetold theoretical and applied insight. This corpus of knowledge began to identify the causes of, and recommend safeguards to, eyewitness misidentification. Legal systems subsequently acknowledged the difficult task of assessing the probative value of eyewitness testimony in court (see Leippe, 1995).

Decades of archival case research has now demonstrated that eyewitnesses can provide honestly believed but mistaken identifications (Loftus, 1979; Loftus & Ketcham, 1991). The consequences of misidentification can be dire; eyewitness misidentification is the highest single factor contributing to wrongful convictions in North America (Association in Defence of the Wrongfully Convicted, 2014; Innocence Project, 2014). As the concerns of memory evidence voiced in the early 20th century have been proven founded, inquiries into wrongful convictions have led to hundreds of exonerations (Huff & Killias, 2013; Scheck, Neufeld & Dwyer, 2000).

Applied research has assisted in unearthing factors that diminish eyewitness identification accuracy (Cutler & Penrod, 1995). Items of interest in eyewitness testimony research are broadly placed into two categories: system and estimator variables (Wells, 1978). *Estimator variables* are factors known to influence identification accuracy that are outside the control of the criminal

justice system. The impact of these variables on actual cases may only be estimated *post hoc* and include witnessing and identification conditions (Wells & Olson, 2003). For example, in *State v. Cromedy* (as cited in Meissner & Brigham, 2000), the New Jersey Supreme Court ruled that, in cases where a critical cross-race identification is made, a judicial caution must be provided to jurors in order to bring to mind the existence of the *cross-race effect*. Sporer (2001) defines the cross-race effect as a robust research finding that refers to a "performance deficit" (p. 170) in recognizing faces of another ethnicity, compared to recognizing faces of one's own ethnicity.

Unlike estimator variables, *system variables* are under the control of the criminal justice system. Such items include all aspects of eyewitness lineup administration, police questioning techniques, and the *retention interval*: the duration of time between the event and subsequent identification (Cutler, Penrod & Martens, 1987). For example, in *Regina v. Maragh* (2003), defence attorneys applied, unsuccessfully, to admit expert testimony on the dangers of requesting an identification from an eyewitness using a *simultaneous lineup* as opposed to a *sequential lineup*. Research has suggested that when a lineup is viewed by an eyewitness sequentially, they are less likely to make an incorrect identification (Lindsay & Wells, 1985; Lindsay, 1999). [1] Research that suggests the need to modify system variables has direct implications on legal policy and procedure. As such, recommendations made from the study of system variables have resulted in international criminal justice policy reform (e.g., Technical Working Group, 1999; Law Reform Commission of Canada & Brooks, 1983).

A sub-field of research on eyewitness testimony, under which the present study may be classified, emphasizes the role of juror as fact-finder. In these studies, participants take on the

[1] Sequential lineups have been adopted in several jurisdictions. However, the assertion that sequential lineups are superior to simultaneous lineups is contested; see Gronlund, Wixted and Mickes (2014).

role of juror in order to evaluate evidence presented in a criminal case. Traditionally, these studies involve eyewitness testimony as the primary or sole evidence against a defendant. A critical component in many of these studies is the inclusion of *procedural safeguards* (PSs). Broadly, PSs are in-court proceedings carried out with the intention of "focusing (juror) attention on the issues surrounding eyewitness testimony" (Katzev & Wishart, 1985, p. 733).[2] While the form and content of PSs vary, they share in common the goal of preventing wrongful convictions based on unreliable eyewitness evidence (Bromby, MacMillan & McKellar, 2007). Specifically, PSs include eyewitness cross-examination, eyewitness evidence *voir dire* hearings, expert testimony, and judicial cautions (Martire & Kemp, 2009). Each PS may address research findings on system and estimator variables into the courtroom to be assessed by a jury. By independently varying conditions within experiments, post-trial evaluations permit the gathering of insight into the perceptions and decision-making process of jurors. In turn, the impact of PSs on juror determinations, such as verdicts, may be empirically evaluated.

The present study evaluated the effect of expert testimony and judicial caution on jurors' perceptions of two recurring issues in cases involving eyewitness identification evidence: eyewitness confidence and photoarray lineups. The literature review begins by focusing on developments within the Canadian legal system on the admissibility of eyewitness expert testimony in Canada. This analysis is supplemented with an overview of the various methodologies used in, and conclusions drawn from, research examining expert testimony as a PS. Attention is then turned to the use of judicial cautions as a PS in the Canadian legal system and introduces the model drafted by the Canadian Judicial Council (2012). Research examining

[2] This conceptualization was provided by Katzev and Wishart (1985) to describe judicial cautions as opposed to all PSs. The word "juror" was placed in parenthesis by the author to broaden the definition.

the influence of judicial cautions on jurors is summarized. Next, the topic of the *confidence-accuracy* (CA) relationship in eyewitness testimony is examined. A review of studies analyzing the effect of eyewitness confidence on juror decision-making is provided. The literature review concludes by surveying research on eyewitness lineups, a component of applied eyewitness testimony research that has provoked considerable attention. Structural and theoretical considerations in the construction, deployment and assessment of eyewitness lineups are presented. In Study 1, two lineups were empirically assessed for fairness in a pilot study. In Study 2, a *minimal trial* (Martire & Kemp, 2009) – a simulated court proceeding conducted for research purposes – was conducted. Jurors completed a ten-item Questionnaire and results were examined to determine how eyewitness evidence and procedural safeguards impacted the perceptions of jurors. Results were discussed in view of current Canadian trial proceedings.

Eyewitness Expert Testimony

Among procedural safeguards that may be presented in court is the testimony of an eyewitness expert. Eyewitness experts serve a context-building function by providing *social framework* testimony (Walker & Monahan, 1987). Where an eyewitness account is presented as evidence, experts may be requested by counsel to attest to the nuances of human memory and discuss the potential impact of case-relevant variables on eyewitness accuracy. Rather than offering an ultimate opinion on the contents of an eyewitness' memory, the jury is educated on general conclusions from social science research to assist in their fact-finding function (Cutler & Kovera, 2010; Leippe, 1995; Malpass, Ross, Meissner, & Marcon, 2009). Whether eyewitness expert testimony is an appropriate procedural safeguard (PS) is contested.

Supporters contend that expert testimony "can help jurors better judge the veracity of eyewitness testimony" (Leippe, Eisenstadt, Rauch, & Seib, 2004, p. 525). In this vein, compelling arguments have been made in support of allowing eyewitness experts into court to serve this pedagogical function (e.g., Fradella, 2006; Leippe, 1995; Wise, Dauphinais & Safer, 2007; Yarmey, 2001). On the contrary, critics argue that expert testimony may "cause jurors to become overly skeptical of eyewitness evidence" (Leippe et al., 2004, p. 525). The probative value of expert testimony clearly must be balanced with its potential prejudicial effect. To meet this end, Canadian courts have developed guidelines surrounding the admissibility and approved scope of eyewitness expert testimony through decades of common law.

Tests of Admissibility

The benchmark United States Supreme Court case of *Daubert v. Merrell Dow Pharmaceuticals* (1993) (hereafter *Daubert*) set out admissibility standards for expert testimony

that would later be mirrored in Canada. In *Daubert*, the Supreme Court held that the Federal Rules of Evidence supersedes the general acceptance standard of expert testimony set out in *Frye v. United States* (1923) (hereafter *Frye*). Under *Frye*, expert opinion testimony was admissible if based on techniques and methodologies generally accepted as reliable by the relevant scientific community. *Daubert* expanded the admissibility standard to a four-tiered test, delineated in the Federal Rules of Evidence (Rule 702; see Federal Evidence Review, 2014). A qualified expert may testify if the testimony: a) includes scientific, technical or other specialized knowledge that will assist the trier of fact to understand the evidence or determine a fact in issue; b) is based on sufficient facts or data; c) is the product of reliable principles and methods; and d) reliably applies the principles and methods to the facts of the case. *Daubert* established that trial judges serve a 'gatekeeper' function to determine that the proposed testimony is scientific and that it will assist the trier of fact. Expert testimony must relate to any issue in the case to be relevant and helpful. Further, the Court advised that federal judges, when considering admissibility, may review whether the proffered information is falsifiable, refutable or testable; whether the theory or technique was peer-reviewed or published; the known or potential error rates; and finally, as per *Frye*, has 'general acceptance' within the relevant scientific community.

The *Daubert* decision in the United States is analogous to the decision in the Canadian case of *Regina v. Mohan* (1994) (hereafter *Mohan*). As summarized by Justice Finlayson, Mohan was a pediatrician that stood accused of four counts of sexual assault on female patients aged 13 to 16. A psychiatrist offered expert testimony for the defence at trial. The psychiatrist contended that the likely offender of the first three assaults would be a paedophile, while the fourth assault indicated the likely offender was a sexual psychopath (para. 8). The trial judge opined that the

expert's experience with sexually assaultive physicians was insufficient to warrant any conclusion and ruled the evidence inadmissible (para. 11). Mohan was convicted on all counts. On appeal, the Ontario Court of Appeal ordered a new trial. In the decision, Justice Finlayson contended that the trial judge had erred in that his conclusions were based on a misapprehension of the evidence (para. 12). The expert's conclusions were not inferred from a select sample of physicians accused of sex crimes, but appropriately drawn from broader psychiatric experience and literature. In essence, the Court of Appeal contended that Justice Bernstein had mistakenly ruled on the sufficiency of the evidence, not the admissibility (para. 12).

In *Mohan*, Justice Sopinka for the Supreme Court of Canada addressed the outstanding issue of the admissibility of expert testimony by establishing a four-tiered test. As in *Daubert*, the expert testimony presented must: a) be logically and legally relevant to the facts of the case; b) necessary to assist the trier of fact; c) not in violation of an exclusionary rule; and d) from a properly qualified expert. Acting judges again were issued a gatekeeping function by being provided the power to decide whether the testimony was relevant to the case. If evidence was determined to be logically relevant, it may still be excluded if its probative value is outweighed by its prejudicial effect, if the value offered in the testimony is not commensurate to the time required to present it, or if it could influence the trier of fact out of proportion to its reliability. Expert evidence should not be admitted if it will be misused or distort the fact-finding process or will confuse the jury. Finally, testimony is only necessary when outside of the experience and knowledge of a judge or jury.

‘Limited Provisions’: The Scope of Eyewitness Expert Testimony

Provisions surrounding the scope of eligible content to be discussed by eyewitness experts were set out in the *voir dire* of expert evidence in *Regina v. Henderson* (2009) (hereafter *Henderson*). In *Henderson*, the accused was charged with first degree murder in a shooting case that relied heavily on four eyewitnesses slated to testify for the Crown. The Crown alleged that the accused and victim were involved in a dispute at a nightclub. Henderson later attended a party armed with a handgun and shot the victim twice at point blank range. Defence attorneys filed a motion to allow Elizabeth Loftus to testify on eyewitness evidence. However, the Crown opposed this motion on the grounds that sufficient instruction could be provided through a judicial caution on the frailties on eyewitness testimony. In the decision, Justice Sinclair acknowledged that jurors may overestimate the reliability and strength of eyewitness testimony and determined Loftus’ testimony would be of desirable probative value. Justice Sinclair applied the *Mohan* test and concluded that Loftus would be eligible to testify, yet subject to limited provisions. Under these limits, Loftus was not permitted to comment on the reliability of any particular witness’ evidence or on any case-relevant factors other than through general comments and hypothetical questions. Loftus was also not permitted to express an opinion on the validity, reliability or bias of the actual photo pack line-up in the case. She was permitted to testify on issues surrounding lineup administration and assessments of fairness. Henderson was convicted of first degree murder and received the mandatory 25 year imprisonment sentence. A subsequent appeal was denied in *Regina v. Henderson* (2012).

A Taxonomy of Eyewitness Expert Studies

The *limited provisions* clause outlined in *Henderson* is identical to the *general testimony* form investigated in research by Fox and Walters (1986). The researchers questioned whether juror perceptions (e.g., verdict) varied across types of testimony (general or specific) and eyewitness confidence levels (low or high). Specific expert testimony involved the discussion of case-specific details, while general testimony limited testimony to social frameworks and broad conclusions akin to provisions outlined in *Henderson*. Compared to a control group viewing no expert testimony, jurors who viewed testimony (general or specific) rendered significantly less guilty verdicts, suggesting that jurors may have disbelieved the eyewitness testimony. In addition, both types of testimony significantly lowered jurors' ratings of their belief that the eyewitness made a correct identification. Few differences were found between the types of testimony. Overall, results suggested that eyewitness expert testimony strongly influences jurors regardless of whether the testimony is general or specific in nature.

McCloskey and Egeth (1983) suggested that there may be a mismatch between the assistance that must be provided to jurors in their evaluations of eyewitness testimony and the actual effects of expert testimony. The test for the appropriateness of expert testimony, they hold, is found in the *discrimination rationale*: "…jurors clearly need help in discriminating accurate from inaccurate eyewitnesses but may not need to be made more skeptical overall" (p. 556). From this postulate grew additional consideration to the consequences of eyewitness expert testimony on jurors. Cutler, Dexter and Penrod (1989) argued that the appropriateness of

eyewitness expert testimony in the courtroom may be determined by examining three plausible effects of its presentation: *juror sensitization*, *juror confusion*, and *juror skepticism*.[3]

Juror sensitization is considered a desirable effect of expert testimony. Cutler et al. (1989) contend that sensitization involves two components, being knowledge (the awareness of units of information) and integration (the formation of a judgment congruent with this information). According to Martire and Kemp (2011), jurors sensitized to eyewitness identification evidence "will be more likely to believe eyewitness identifications made under relatively good witnessing conditions, and less likely to believe eyewitness identifications made under relatively poor witnessing conditions" (p. 25-26). In experimental studies, testing for sensitization requires the independent variation of both 'good' and 'poor' witnessing and identification conditions. For example, Devenport, Stinson, Cutler and Kravitz (2002) found evidence for juror sensitization by expert testimony to lineup instruction bias, through decreased ratings of procedural fairness (but only when foil bias was absent). Evidence for sensitization is found with significant interactions between varied witnessing or identification conditions and exposure to expert testimony. Sensitization is evident where jurors render ratings on a dependent variable in a manner consistent with the information provided by the eyewitness expert.

Juror confusion arises where jurors "misinterpret, overgeneralize, or misapply the information presented by the psychologist, and so may come to unwarranted conclusions" (McCloskey, Egeth, & McKenna, 1986 as cited in Cutler et al., 1989, p. 312). Martire and Kemp (2011) define juror confusion as "a situation where jurors are sensitive to the manipulation of witnessing and identification conditions, but respond to them in a manner contrary to the experts'

[3] Though this review refers to these effects as resulting from exposure to expert testimony, these effects may be produced by other PSs (i.e., judicial caution, cross-examination) and interpreted identically.

advice" (p. 31). For example, a study by Lindsay (1994) found evidence for juror confusion where testimony by an expert produced a significant increase in jurors' perceived fairness of a biased lineup. As with sensitization, evidence for juror confusion is found when significant interactions are obtained on a dependent variable. However, unlike sensitization, jurors' ratings are in a manner contradicting the information provided by an expert.

Juror skepticism results when jurors undervalue, disbelieve, or distrust eyewitness testimony. Martire and Kemp (2011) note that skepticism results when an eyewitness expert undermines the credibility of eyewitness testimony and "therefore does little to help the trier of fact discriminate between likely accurate and likely inaccurate identifications" (p. 26). Evidence for juror skepticism is traditionally measured by examining main effects of expert testimony on a dependent variable, such as verdict. For example, Geisleman et al. (2002) (as cited in Martire & Kemp, 2011) found evidence for skepticism where expert testimony produced a significant increase of not guilty verdicts. Cutler et al. (1989) note that juror skepticism may be argued to be desirable, noting the poor accuracy rates (~50%) stemming from realistic field experiments. However, skepticism may also be interpreted as an indication of the expert usurping the fact-finding function of the jury.

Beyond the plausible effects of expert testimony on jurors, Martire (2008) and Martire and Kemp (2009/2011) outlined a hierarchically organized taxonomy of studies that are classified in view of the total knowledge value obtainable from each design. *Sensitivity to eyewitness accuracy* (SEA) designs permit the testing of juror discrimination by exposing jurors to authentic eyewitness testimony. These designs involve staging a crime for eyewitnesses and presenting authentic eyewitness testimony to jurors. This approach allows comparisons between

the "objective accuracy of the eyewitness identification with juror evaluations of the accuracy of that identification" (Martire & Kemp, 2011, p. 233). While these studies (e.g., Martire & Kemp, 2009; Wells et al. 1980; Wells & Wright as cited in Wells, 1986) provide the notable benefit of being able to determine whether expert testimony helps jurors discriminate accurate from inaccurate eyewitnesses, conclusions drawn from SEA studies done to date are limited to the staged crimes presented in the studies.

Sensitivity to expert opinion (SEO) designs expose jurors to fictional eyewitness testimony transcripts. In SEO designs, witnessing and identification conditions vary as either 'poor' or 'good' across exposure to expert testimony. While researchers are unable to determine whether jurors are made better able to discriminate accurate from inaccurate eyewitnesses, this approach permits researchers to ascertain whether jurors are sensitized, confused, or made skeptical towards eyewitness testimony. By collecting verdicts and appraisals of defence case strength, prosecution case strength, and related measures, researchers may systematically investigate if and how expert testimony alters the weight afforded by jurors across factors "identified as correlates of eyewitness accuracy" (Martire & Kemp, 2009, p. 226).

SEO designs are superior to *response to expert evidence* (REE) designs, which hold witnessing and identification conditions static. Researchers manipulate only whether jurors are exposed or not exposed to expert testimony, and therefore may only draw conclusions as to whether expert testimony influences juror decision making, or not. A substantial shortcoming of these studies, noted by Cutler, Penrod, and Dexter (1989), is that researchers confound sensitivity and skepticism by failing to directly vary witnessing and identification conditions (for example, Fox & Walters, 1986; Hosch, Beck & McIntyre, 1980; Maas, Brigham & West, 1985).

These designs leave researchers unable to determine whether changes in a dependent variable are warranted (Martire & Kemp, 2009).

Usurping the Jury?

Of the sixteen SEO studies reported in a literature review by Martire and Kemp (2011), eyewitness expert testimony most often resulted in producing only juror skepticism (7; 43.75%). Skepticism was predominately evident in studies through expert testimony producing a significant main effect of increasing not guilty verdicts. Only three studies (18.75%) found evidence that expert testimony produces juror sensitivity. For example, juror sensitivity was evident where jurors who had viewed expert testimony were more likely than jurors not exposed to expert testimony to believe an identification made under 'good' conditions as opposed to 'poor' conditions.[4] Two studies (12.5%) produced a combination of skepticism and sensitivity. The remaining studies produced confusion (2; 12.5%), combined skepticism and overbelief (1; 5.75%) or no effect (1; 5.75%).

Overall, there is strong evidence from SEO studies that eyewitness expert testimony is highly influential, but may produce less than ideal outcomes. Regardless, scholars have aggressively advocated for a place in court for the eyewitness expert. Including expert testimony in the present study permits conclusions to be drawn as to its efficacy as a PS and permits a direct comparison of the effects of expert testimony and judicial caution.

[4] Jurors may be already sensitive to witnessing and identification conditions. This is evident when there is a significant main effect of a varied case condition, in the appropriate direction, on a dependent variable.

Judicial Cautions of Eyewitness Identification Evidence

With the "ever-present risk of a miscarriage of justice" (*Regina v. Turner*, 2012, para. 110), courts recognize the need for procedural safeguards (PSs) in circumstances of trial-by-jury. However, the four-tiered *Mohan* test of admissibility has proven problematic for attorneys seeking to submit expert testimonial evidence. Numerous Canadian cases have developed a precedent for judges to opt for the use of a judicial caution: "instructions or warnings to a jury in criminal trials (that) arguably can provide a safeguard against erroneous convictions based upon unreliable eyewitness evidence" (Bromby et al., 2007, p. 305). This review will examine developments within Canadian law and findings from international research to illuminate the issues pertinent to understanding the current form of a model caution drafted by the Canadian Judicial Council (2012).

The Caution as Standard Practice

Perhaps no other case decision better illustrates the preference in Canada to use a judicial caution instead of an eyewitness expert than that found in *Regina v. McIntosh and McCarthy* (1997) (hereafter *McIntosh*). In *McIntosh*, two accused appealed convictions of aggravated assault and nine other offences from indictments surrounding a robbery at a dry cleaning store. The appellants held the original trial judge had erred in refusing to allow expert testimony from psychologist Daniel Yarmey. Yarmey was requested by the defence to testify on several factors pertaining to eyewitness testimony, including the cross-race effect, the forgetting curve, post-event information and photoarray lineups. Justice Finlayson upheld the trial judge's decision to reject the expert testimony, and proceeded to characterize judicial cautions on the frailties of identification evidence as more than adequate as safeguards (para. 22). The proper incorporation

of eyewitness literature into the courtroom, Justice Finlayson contended, should be in updating the caution; not introduced as testimony. Justice Finlayson was loquacious in his pointed skepticism toward whether eyewitness research was constitutive of an accepted body of scientific research worthy of opinion testimony.

These latter sentiments were dismissed in a decision by Justice MacKenzie in *Regina v. Myrie* (2003), who acknowledged that eyewitness memory research is a recognized body of scientific knowledge. As in *McIntosh*, the key issue in this *voir dire* was the admissibility of expert evidence as offered by Daniel Yarmey. In his decision, Justice MacKenzie held concern over the "substantial and unacceptable risk of the expert's opinion usurping the function of the jury as fact-finder" (para. 32) and dismissed the application. Following *Mohan*, Justice MacKenzie opined that that proffered expert opinion evidence is not outside the knowledge of the judge or jury. Justice MacKenzie quoted Justice Finlayson in *McIntosh* to restore the Court's opinion that expert testimony on the impact on memory from brief exposure time and stress are not items "outside of the normal experience of the triers of fact" (para. 27). The proposed evidence, therefore, failed to meet the requirement of necessity under *Mohan*. A further concern lied in the irrefutability of the generalizations made by eyewitness experts. Quoting from *Regina v. D.D.* (2000), Justice Finlayson characterized eyewitness expert testimony as not lending itself "to future advancements in knowledge and understanding" (para. 66). Judicial cautions were awarded preference to expert testimony. Applications for expert testimony by Yarmey on photoarray lineup evidence were similarly denied in *Regina v. Eakin* (1994) and *Regina v. Nyugen* (1997).

Judicial cautions do not assume a uniform and singular form across cases. Trial judges in Canada may produce their own caution to meet the unique needs and facts of a case, at their discretion and with or without consultation. The content of these cautions are subject to appeal if they include plausible errors in the direction provided to the jury. This consideration was at issue in *Regina v. Hay* (2013) (hereafter *Hay*). In *Hay*, the appellant contended that a judgment from the Ontario Court of Appeal affirming his conviction of murder and attempted murder stemming from a shooting in a Toronto nightclub was in error. At issue was the wording of a caution used by the trial judge on the frailties of eyewitness testimony. The final paragraph of the caution stated that the jury was "entitled to convict on the evidence of even a single eyewitness" (para. 50). Justice Rothstein, for the majority, opined that the caution issued to the jury by the judge, as a whole, did not err in instructing the jury in this manner. The judge at trial instructed the jury to look for additional evidence confirming the identification and urged jurors to approach problematic identification evidence provided by one eyewitness with greater caution than other, more reliable witnesses. In the majority view, these mentions absolved the point of warranting a new trial.[5] Notwithstanding the frailties of eyewitness identification, *Hay* withheld that juries may enter a conviction "even where the Crown has relied on only a single eyewitness" (para. 40). In dissent, Justice Fish contended that instructing the jury that they may convict on the problematic evidence of the single eyewitness was a fatal error.

Research on Judicial Cautions

Jury instructions drafted in *U.S. v. Telfaire* (1972) (hereafter *Telfaire*) became the most widely used standardized caution on eyewitness testimony in the United States (Woocher as

[5] A new trial was ordered due to forensic hair evidence that was presented as a second, unrelated ground for appeal.

cited in Greene, 1988). In *Telfaire*, the United States Supreme Court affirmed the conviction of robbery and sentence under the Federal Youth Corrections Act (1950) in a case that was based on uncorroborated testimony of a single eyewitness of a different race of the accused. The decision of the Court put forward a script of special instructions to be stated by a judge to a jury in cases involving eyewitness identifications. The instructions directed the jury to consider pertinent factors involved in eyewitness identification, including the opportunity the witness had to observe the offender, viewing conditions, and the composition of the lineup. The *Telfaire* instructions prompted researchers to examine the effects of eyewitness testimony judicial cautions on juror decision-making.

Results from a minimal trial conducted by Greene (1988) suggested that the *Telfaire* instructions primarily produced a dramatic decrease of guilty verdicts from jurors who were exposed to the caution. Scores on a multiple-choice questionnaire demonstrated that jurors exposed to *Telfaire* instructions were made no more knowledgeable on eyewitness identification factors than that of a control group. A second experiment demonstrated that simplifying the instructions improved juror comprehension of factors to consider in the evaluation of eyewitness testimony. Similar attempts at simplifying judicial cautions by Bornstein and Hamm (2012) have produced some improvements in subjective comprehension of jurors without a significant effect on verdicts.

Cutler et al. (1990) re-examined the effect produced by the *Telfaire* instructions on juror decision-making. In this design, the researchers independently varied multiple witnessing and identification conditions (disguise, weapon focus, retention interval, and instruction bias) in a fictional trial. The researchers found the *Telfaire* instructions “completely ineffective” (p. 1205)

in sensitizing jurors to eyewitness evidence and recommended the development of instructions that do sensitize jurors. Both Greene (1988) and Cutler et al. (1990) laid out strong arguments for the inefficiency of the *Telfaire* instructions. In general, the researchers contended that the instructions did not address findings from eyewitness research nor did they provide information on numerous variables known to influence memory.

Other investigations of judicial cautions have produced mixed results. Paterson, Anderson and Kemp (2013) found some evidence that a specific judicial warning regarding dangers of co-witness discussion induces sensitization with an absence of skepticism. O'Connor (2013) examined whether the New Jersey judicial eyewitness caution sensitized jurors to the cross-race effect. The instructions did not influence verdicts, but jurors rated the New Jersey caution more informative and clearer than a basic warning. Katzev and Wishart (1985) exposed jurors to just a judicial caution, a judicial caution with a case summation, or a judicial caution, case summation, and commentary on eyewitness testimony. The researchers found that jurors who viewed all three safeguards rendered the lowest rates of pre-deliberation guilty verdicts, took less time to deliberate, and rendered the least individual guilty verdicts compared to groups that had viewed just the caution or the caution and the summation.

A Standardized Canadian Model

Justice Lysyk in *Regina v. Savoy* (2000) noted the "many examples of judicial cautions concerning the care that must be exercised with respect to reliance upon a witness's identification" (para. 29). Perhaps in response, the Canadian Judicial Council (2012) formalized a model jury instruction for use in cases that rely primarily on eyewitness evidence that is challenged by the defence. The caution directs the jury to consider several factors under four

broad categories: 1) the reliability of the eyewitness; 2) the eyewitness' observation circumstances; 3) the description given by the eyewitness; and 4) the procedure followed for the identification. The caution stresses the potentially misleading nature of eyewitness confidence: "Eyewitness identification may seem more reliable than it actually is because it comes from a credible and convincing witness who honestly but mistakenly believes that the accused person is the one he or she saw committing the offence" and continues, "there is little connection between great confidence of the witness and the accuracy of the identification. Even a very confident witness may be honestly mistaken. A very confident witness may be entirely wrong with respect to his or her identification evidence" (p. 60).The caution also directs jurors to appraise the fairness of a lineup: "Did the other participants in the line-up share the physical characteristics of the accused? Were the photos similar?" (p. 62). The caution includes portions in parenthesis that may be included, omitted, or modified by the acting judge, depending on facts and circumstances of the case at hand.

A fundamental difference between the caution drafted by the Canadian Judicial Council and that drafted by the United States Supreme Court is that the Canadian model incorporates more empirical findings from eyewitness research. Further, the Canadian model places considerably more emphasis, in length and wording, on factors influencing identification accuracy. The inclusion of examples and discretionary optional sections in the Canadian model, wherein the presiding judge may address case-specific facts, is a notable advantage to the static *Telfaire* model. To date, only two studies (Cutler et al., 1990; Martire & Kemp, 2009) have directly compared the effects of a judicial caution to expert testimony on the perceptions of jurors, and no studies have examined the Canadian model. Having been drafted by a federal

body, the Canadian model stands likely to become the most widely-used caution in Canada. Given the high likelihood that this caution will be used in court, it is beneficial to understand how these instructions may affect the perceptions of jurors under certain case circumstances. This evaluation encompasses these important developments.

Juror Decision-Making: Putting Confidence into Context

The confidence of an eyewitness in an identification decision is the most regularly reported *postdictor* of accuracy in the eyewitness literature (Smith, Lindsay, & Pryke, 2000). Postdictors are variables relating to an identification decision (e.g., response latency), collected to ascertain what can usefully discriminate accurate from inaccurate eyewitnesses (see for example Charman & Cahill, 2012; Sauerland & Sporer, 2009). The collection of decision confidence is but an extension of classic experiments that test the ability of a learner to assess the quality of encoded stimuli. Early memory researchers asked learners to estimate their ability to recall the necessary information "at the time of the memory test" (King, Zechmeister, & Shaughnessy, 1980, p. 330). Related research reviewed in a study by Hart (1965, as cited in Schwartz, Benjamin & Bjork, 1997) found that correct answers were acutely associated with a person's "feeling of knowing" (p. 134). This tenet is a basic principle under *trace access theory*, which has been highly influential on eyewitness identification literature.

Trace access theory presumes that memory traces are created from events and experiences. These traces vary in strength, with stronger traces better supporting accurate performance on a memory task than weaker traces (Roediger, Wixted, & Desoto, 2012). Under trace access theory, an eyewitness's self-reported confidence in an identification decision is believed to be based on an internal assessment of the salience of the memory (Busey, Tunnicliff, Loftus & Loftus, 2000). This connection appears rational and has been propagated and legitimized in the United States Supreme Court.[6] However, researchers have noted several issues with this generalization.

[6] *Neil v. Biggers* (1972)

As noted by Brewer, Weber and Stemmler (2007), most eyewitness researchers characterize confidence as a generally unreliable indicator of identification accuracy. Eighty-seven percent of expert researcher respondents to a survey by Kassin et al. (2001) considered the general conclusion that "an eyewitness's confidence is not a good predictor of his or her identification accuracy" (p. 408) is a finding reliable enough to testify about in court. Brewer et al. (2007) posit that, in general, laboratory studies have produced results suggesting that inaccurate eyewitnesses are frequently confident. Similarly, actual cases have demonstrated that an absolutely positive eyewitness can be entirely incorrect (Thompson-Cannino, Cotton, & Torneo, 2009).

A meta-analysis of the *confidence-accuracy* (CA) relationship by Bothwell, Deffenbacher and Brigham (1987) (as cited in Penrod & Cutler, 1995) found an average correlation of $r = .25$ across 35 studies. Wells and Murray (1984) found only an average $r = .07$. Findings from Sporer, Penrod, Read and Cutler (1995) similarly supported the conclusion that the overall CA relationship is relatively weak ($r = .23$). However, correlations may oversimplify the nature of the CA relationship. First, an important moderator noted by Sporer et al. (1995) is that eyewitnesses who make a decision, or *choosers*, are more likely to be accurate than *non-choosers*. Decisions by choosers (in isolation) generally have stronger CA correlations. Second, researchers have rejected the correlation approach in favour of a calibration approach.

According to Juslin, Olsson and Winman (1996), CA assessment under calibration requires the eyewitness to provide the "subjective probability that the identified person is identical to the culprit" (p. 1305) on a scale (0%, 10%, 20%…100%). These subjective probabilities (decision confidence) are then compared to their "corresponding relative

frequencies of correct identifications within the confidence categories" (p. 1305). This form of analysis allows information to be gained regarding the association between subjective confidence and objective probability of a correct identification. Under this design, researchers have concluded that point-biserial correlation analyses have underestimated the CA relationship (Olsson, Juslin & Winman, 1998; Weber & Brewer, 2003; Brewer, Keast & Rishworth, 2002). However, researchers continue to find patterns of under- and over-confidence (e.g. Juslin, Olsson & Winman, 1996; Wells & Brewer, 2006). For example, Wells and Brewer (2006) found that 90 – 100% confidence levels were associated with lower accuracy (75% and 90%). The researchers emphasized that, although calibration offers a more optimistic characterization of the CA relationship, results in no way challenge the overall expert consensus. As such, experts warn jurors that an eyewitness' high confidence does not necessarily amount to accuracy.

Confidence Malleability

Questions of the reliability of a confidence declaration may also be raised in light of numerous investigations that have demonstrated that confidence is susceptible to inflation from environmental variables. *Postidentification feedback* refers to "information provided by a lineup administrator to an eyewitness immediately following the selection of a suspect" (Terrance, Thayer & Kehn, 2006, p. 76). There has been robust support that postidentification feedback increases confidence in an identification from inaccurate eyewitnesses. For example, Wells and Bradfield (1998) found that confirming feedback ("Good, you identified the actual suspect") significantly inflated retrospective judgments of certainty, quality of view, clarity of their memory and speed of identification. Luus and Wells (1994) found that informing witnesses that other witnesses had selected the same suspect produced similar retrospective confidence

inflation. Wells, Ferguson and Lindsay (1981) found support for their hypothesis that the pretrial brief given by an attorney to a witness increases the witness' thought about the decision which, in turn, increases confidence.

Confidence is vulnerable to inflation whether cues are explicit or implicit. Garrioch and Brimacombe (2001) instructed lineup administrators to avoid providing witnesses with any feedback regarding their identification. However, lineup administrators who were led to believe a certain suspect in a target-absent lineup was guilty subtly and unintentionally conveyed confirming feedback (i.e., eye contact, vocal intonation) that significantly inflated the confidence of the eyewitnesses. Smalarz and Wells (2013) found that evaluators believed mistaken witnesses that had received post-identification feedback at levels nearly identical to the accurate eyewitnesses. These findings have lent empirical support to the recommendation of *double-blind lineup administration* (Wells, 1988; Wells et al., 1998).[7] When lineup administrators collect a confidence assessment immediately following a positive identification, the eyewitness is not exposed to any influences that may artificially inflate confidence. Indeed, when eyewitnesses make a quick selection from a lineup accompanied with a high confidence declaration, this can be considered a useful (though not absolute) determinant of accuracy (Brewer & Wells, 2006).

Ultimately, however, mistaken eyewitnesses confident in their decisions can and do play a pivotal role in producing a wrongful conviction. Jurors who are sensitive to the numerous factors to consider when evaluating testimony would abstain from relying exclusively on confidence as a lone determinant of culpability. There is substantial evidence, however, that such

[7] In Canada, a recommendation of the double-blind procedure was made in the Inquiry Regarding Thomas Sophonow (Cory, 2011); a publication from the Province of Manitoba made in response to a wrongful murder conviction due in part to eyewitness misidentification.

considerations are beyond the ken of the jury. While 95% of experts responding to a survey by Kassin et al. (2001) agreed that confidence was malleable, only 50% of laypersons responding to a survey by Benton, Ross, Bradshaw, Thomas and Bradshaw (2006) agreed. Similarly, only 38% of respondents to Benton et al. (2006) agreed about the nature of the CA relationship, compared to 87% of research experts in Kassin et al. (2001). Elderly citizens, judges, criminal lawyers and police officers have rendered opinions at odds with those of experts' on factors including time estimations of events, reliability of confidence, and the cross-race effect (Yarmey & Jones, 1982). With this apparent discrepancy in beliefs between prospective jurors and experts, it follows that researchers have examined the pertinent issue to this study: the effect of eyewitness confidence on the perceptions of jurors.

Eyewitness Confidence and Jurors

Numerous studies have investigated the relationship between eyewitness confidence and juror perceptions. Wells, Lindsay and Ferguson (1979) designed an experiment exposing jurors to the examination of eyewitnesses who had viewed a staged crime. The researchers found that eyewitness accuracy and juror belief of the eyewitness was independent; jurors were unable to discriminate between inaccurate and accurate eyewitnesses. Only in a condition of leading cross-examination were jurors more likely to believe an accurate eyewitness. Additionally, jurors' ratings of witness confidence accounted for nearly 50% of the variance in jurors' belief in the testimony. Whitley and Greenberg (1986) similarly found that increased eyewitness confidence increased jurors' perceived accuracy of the witness' account of a robbery, suspect description, and identification decision. Though lower than Wells et al.'s (1979) study, confidence accounted for a significant 7% of variance in perceived credibility. Lindsay, Wells and Rumpel (1981)

revisited this paradigm by first exposing eyewitnesses to a staged theft designed to yield low- (33%), moderate- (50%), or high- (74%) identification accuracy rates. Eyewitness accuracy produced no effect on jurors' decisions to believe the witness. Eyewitness confidence, however, produced a significant main effect on jurors' decisions to believe the witness. Although statistically significant, the researchers concluded that the differences observed in witness confidence ratings across conditions whether accurate (4.3, 4.3 and 5.0) or inaccurate (3.4, 3.8 and 3.9) were negligible. Overall, jurors overbelieved eyewitnesses where accuracy is low and confidence is high. Beaudry et al. (2013) found that perceived eyewitness confidence was positively correlated with jurors' decisions to believe the eyewitness. However, only 53.24% of jurors correctly believed identification decisions from eyewitnesses.

Wells et al. (1981) found that briefing inaccurate eyewitnesses regarding potential cross-examination questions significantly inflated retrospective confidence ratings. In turn, briefing (but not accuracy) significantly predicted juror verdicts. Remarkably, 72% of jurors believed that briefed but inaccurate eyewitnesses made an accurate identification, while a substantial 44% believed not briefed but inaccurate eyewitnesses. Only 40% of jurors believed not briefed, accurate eyewitnesses, and 45% believed accurate, briefed eyewitnesses. Overall, the confidence of an eyewitness is susceptible to external influence. Jurors should not conclude that the high confidence of an identification decision reported by eyewitnesses is deterministic of accuracy.

Cutler, Penrod and Stuve (1988) noted that actual jury decision-making is contingent upon multiple determinants and case factors. The substantial variance accounted for by eyewitness confidence in jurors' decision to believe the witness in Wells et al. (1979) may have been an overestimation stemming from the low number of predictors employed (Kenny, 1985 as

cited in Cutler et al., 1988). The researchers utilized a videotaped minimal trial experiment employing multiple witnesses (two for the defence, two for the prosecution) applying variation of ten witnessing and identification conditions. Only witness confidence had a statistically significant effect on perceived probability that the identification was correct. In total, 54% of jurors in the high confidence condition voted to convict, compared to the 39% in the low confidence condition. The researchers concluded that jurors did not demonstrate a capacity to integrate evidence that suggested a less reliable identification in their decision making. Only confidence significantly influenced the perceptions of jurors.

In summary, the studies reviewed here cumulatively suggest that the fact-finding function of the jury may be undermined by an overreliance on eyewitness confidence. Without an application of evidence-based eyewitness literature in the courtroom, the administration of justice may be said to be in disrepute. At best, the naïve juror is piecemeal in a departure from procedural justice; at worst, they condemn an innocent to incarceration. In Canada, judicial cautions and eyewitness expert testimony are procedural safeguards (PSs) that have garnered much use in legislative arenas. However, there is a virtual absence of any empirical evaluation of these safeguards in their specific form. Perhaps the most apropos eyewitness evidentiary issues argued in Canadian jurisprudence are that of lineup bias and eyewitness confidence. By independently manipulating these conditions across PSs, the influence of these variables on juror perceptions can be established. The following analysis traces the development of guidelines around police eyewitness lineup construction and presentation in Canada.

Eyewitness Lineups in Practice and Research

In *Regina v. Osborne* (2012) (hereafter *Osborne*), the accused had pleaded not guilty to an array of sexual crimes, including sexual assault, stemming from an incident in Toronto. In 2009, P.J., an 8 year-old girl, attended the garbage chute in her apartment building unaccompanied. Confronted upon arrival, P.J. alleged that the Osborne had "asked to see her underwear, attempted to lift her dress and pull her dress down, asked her to touch his penis and then masturbated…" (para. 2). One year after the incident, a police investigation was launched and P.J. selected the accused out of a *photoarray lineup*: an identification test in which eyewitnesses are presented with a set of photos containing a suspect's photo (Cutler & Kovera, 2010). In the Court's decision, Justice Forestell commended the detective for an "above reproach" lineup presentation (para. 42). However, Justice Forestell took issue with the composition of members in the lineup, as the accused was "the only man in the lineup with a shaved head. The one distinctive feature described by P.J. in her original statement…" (para. 44). Further, the Court found that the detective had reassured P.J. that she had made a correct identification, enhancing her level of confidence (para. 49). These errors contributed to the establishment of reasonable doubt and the accused being found not guilty on all counts.

Persons charged with an offense in Ontario may request, without cost, that prosecuting attorneys and police provide all information in their possession that is relevant to the charge(s). This process of disclosure is necessary to uphold the right of full answer and defence, recognized under the Criminal Code of Canada in s. 802 (1) and protected under s. 7 of the Canadian Charter of Rights and Freedoms (life, liberty and security of the person) and under s. 11(d) (right to a fair trial). A decision by the Supreme Court of Canada in *Regina v. Stinchcombe* (1991) (hereafter

Stinchcombe) concluded that the Crown holds a constitutional duty to disclose all relevant information pertaining to a case, regardless of whether it will be introduced at trial. The disclosure of evidence is therefore not limited by its inculpatory or exculpatory nature (Stuart, 2001). All materials used in the administration of an eyewitness lineup, as well as testimony regarding identification decisions, are subject to this disclosure process. Due to this mandate, lineups are presented as evidence in Canadian courts to be evaluated by a jury.

As a system variable, the composition and presentation of eyewitness lineups are under the control of the criminal justice system. When assembling a lineup, an administrator places an actual suspect amongst *foils*: known-innocent, non-suspect persons such as incarcerated prisoners or fellow police officers (Lindsay & Wells, 1980). The vast majority of lineups in North America are conducted using a photoarray, as they are economical in terms of effort, time and cost expenditure when compared to live or videotaped lineups (Turtle, Lindsay & Wells, 2003; Wogalter, Malpass & McQuiston, 2004). Failing to abide by lineup construction guidelines adopted in many jurisdictions can have direct consequences on judicial decisions. As illustrated in *Osborne*, the Canadian judiciary may heavily scrutinize structural and procedural lineup conditions that may bias the eyewitness to select the defendant (see Wells et al., 1998).

Guidelines for lineup administration have been built around literature suggesting that certain conditions can influence rates of false and accurate identifications. For example, *instruction bias* exists where a lineup administrator fails to inform the witness that the perpetrator may not be present in a lineup, suggests that the perpetrator is present, or conveys that the witness is expected to make a choice (Stinson, Devenport, Cutler & Kravitz, 1997). *Presentation bias* includes the use of *simultaneous* (all-at-once) as opposed to *sequential* (one-at-

a-time) lineup presentation (Devenport, Penrod & Cutler, 1997). *Investigator bias* refers to a demand characteristic: the implicit or explicit behaviour by the lineup administrator that influences identification decisions (Fanselow, 1975; Wells & Luus, 1990). *Foil bias* exists where few lineup members match the eyewitness's description of a suspect (Wells, 1993). Experimental methods exist to measure how these biases influence eyewitness identification decisions.

The *lineup-as-experiment* analogy is well-suited to the hypothesis-testing scenario of determining guilt of a suspect (Wells & Luus, 1990). For example, an administrator may hypothesize that a suspect is the offender and employ confederates as foils. Various conditions of the lineup may be manipulated, including the number of foils, appearance of foils, presence or absence of suspect, pre-identification instructions, or mode of presentation. The outcome of such a scenario may be the identification of a suspect, identification of a foil, or a lineup rejection. In such experiments, researchers are often interested in measuring *lineup fairness*: the degree to which a lineup is without bias towards selection of the suspect. The appearance of foils, specifically, is a structural consideration that may be evaluated.

Doob and Kirshenbaum (1973) pioneered an experimental method to measure lineup fairness coined the *mock-witness procedure*. In the mock-witness procedure, lineup administrators provide naïve participants a verbal or written physical description of the suspect, expose them to an eyewitness lineup, and ask that they make a selection (Malpass & Lindsay, 1999). The central assumption of the procedure is logically sound: if a suspect receives selections significantly exceeding that which can be expected from chance alone, then the lineup is biased toward the suspect (Malpass, Tredoux and McQuiston-Surrett, 2007). The null hypothesis for the mock-witness procedure is that the lineup is fair. In order to create a fair

lineup, administrators may aggregate identification decisions from several *mock-witnesses* in testing this hypothesis. In controlled research settings, the misidentification of a suspect that is physically similar to the offender from a target-absent lineup is a *false alarm*. In police investigations, this scenario creates evidence of guilt against an "innocent suspect" (Steblay, Dysart, Fulero & Lindsay, 2001, p. 464).

The ratio of accurate to inaccurate eyewitness identification decisions made under particular conditions has been termed the *diagnosticity ratio*, an index that utilizes a Bayesian approach (Wells & Lindsay, 1980). Using this index, identification decisions obtained using biased lineup procedures, including foil-biased photoarray lineups, have been demonstrated to be less informative of a suspect's guilt than decisions obtained using unbiased lineup procedures (Fitzgerald, Price, Oriet, & Charman, 2013; Lindsay et al., 1991; Malpass & Devine, 1981). As mistaken identification is the leading contributing cause of wrongful convictions, ensuring that foils appear suitably similar to a suspect serves as a safeguard against the danger of misidentifications as a result of deduction (Cutler & Penrod, 1995). Research experts generally agree eyewitnesses are more likely to be accurate when more members of a lineup resemble the suspect (Kassin, Ellsworth & Smith, 1989; Kassin, Tubb, Hosch & Memon, 2001).[8] Evaluations of lineup fairness should be conducted by lineup administrators during construction. If this process is overlooked, defense attorneys may rightfully challenge the evidence "as an issue of justice" (Brigham, Meissner, and Wasserman, 1999, p. S73).

[8] In the 1989 study, 77% of experts surveyed agreed this finding is reliable enough to testify in court. In the 2001 study, 70% agreed.

Lineup Parameter Indices

Several formulations have been put forward to quantify lineup fairness, which assists in research replication and comparison (Malpass et al., 2007). Two concepts are of central importance in evaluating lineup fairness: lineup bias and lineup size. *Lineup bias* is the extent that the suspect is distinctive from the other lineup members (Brigham et al., 1999). *Lineup size* can best be understood by distinguishing nominal size from effective size. The *nominal size* of a lineup is simply the number of persons in a lineup. In practice, nominal size can vary greatly between jurisdictions (Turtle et al., 2003). A parameter of greater importance is the lineup's *effective size*: the number of foils that are plausible alternatives to a suspect. In evaluating plausible alternatives, lineup members that receive no selections from mock-witnesses may be considered null (ineffective) foils.

Lineup Bias. The Lineup Bias index (Doob & Kirshenbaum, 1973) evaluates summary data acquired from the mock-witness procedure to determine if a suspect draws identifications significantly greater than chance, or 1/*k*, where *k* is the nominal size of the lineup. Doob and Kirshenbaum (1973) conducted a *z*-test of proportions on mock-witness selections on a lineup featured in a Canadian case. Mock-witnesses chose the suspect at rates significantly greater than chance, and the researchers concluded the lineup was biased. However, the normal distribution can yield unreliable estimates of the significance of the binomial distribution for rare events where the probability of success is small (Dunning, 1993). Moreover, gathering a large sample size may not be a feasible task for law enforcement personnel looking to evaluate a lineup. Tredoux (1998, 1999) recommends an exact probability approach that may be used in conjunction with normal curve estimates.

As detailed by Tredoux (1998, 1999), lineup evaluators may assume independent, random mock-witness choice. Mock-witness selections may be thought of as a Bernoulli trial (with outcomes of either success or failure) with a fixed probability of success (i.e., correct identification) equalling chance (1/*k*). The practical implication of the assumption of equiprobability of selection under the null hypothesis (i.e., the lineup is fair) entails the presence of effective foils. The cumulative distribution of the number of trials in which a successful choice is made may take on the cumulative binomial probability distribution.[9] The calculated values may be compared to conventional levels of significance to evaluate if suspect selections were significantly greater than that which can be expected by chance.

Effective Size. Malpass (1981) proposed the Effective Size (ES) index, which "represents the degree to which the lineup presents to mock witnesses fewer effective choice alternatives than the nominal size of the lineup" (p. 301). This index was refined by Tredoux (1998, 1999). Unlike the initial ES index, Tredoux' ES considers all suspect selections made by mock-witnesses. The index is bounded at the upper end with *k* (the nominal size of the lineup) and at the lower end with *j* (a minimum of 1):

$$ES = k - \sum_{j=1}^{k} \frac{|O_i - e|}{2_e} \quad (1)$$

Where O_i is the observed number of mock-witness choices that chose lineup member *i* and *e* is the number of alternatives in the lineup. However, as a point estimate, ES does not lend itself to

[9] See Tredoux (1998) for formulas.

inferential reasoning. To account for inherent sampling variation, an alternative but related index may be used, *E* (Tredoux, 1998, 1999): [10]

$$E = \frac{1}{1 - I} \tag{2}$$

Where *I* (the *Index of Diversity*), represents "the probability that two individuals selected at random from the population would be in different categories" (Agresti & Agresti, 1978, p. 206). Evidence that a lineup member insufficiently meets the suspect description is found when fewer mock-witness selections than that of chance are obtained for a particular suspect. In mock-witness terms, Tredoux' *E* "reduce(s) the nominal size of the lineup in proportion with the deviation of the observed frequencies from expectation" (Tredoux, 1998, p. 225). The values of *I* vary between 0 and $(k - 1)/k$. In essence, $I = 0$ represents a scenario where all mock-witnesses choose the same lineup member. An equal distribution of mock-witness selections over all members of a lineup results in a maximum value. Tredoux (1998, 1999) has demonstrated that confidence intervals may be built around *I* to gather upper and lower bounds of a lineup's effective size with *E*.[11]

This inferential approach is desirable to quantify the discrepancy between observed and expected frequencies (Malpass & Lindsay, 1999). In a perfectly fair lineup, the nominal size would be equal to its effective size. An outcome suggestive of lineup fairness would be the containment of the nominal size of the lineup within the 95% CI of ES, with a minimum of six (suspect included) effective members (noted by Wells, 2001). Measurements of lineup fairness reviewed here have been endorsed by researchers assessing lineups used as stimuli in

[10] Although *ES* and *E* are independent indices, Malpass et al. (2007) demonstrated that the two are strongly correlated across wide ranges of lineup sizes and mock-witnesses, $r = .97$ - $.99$, all ps < 0.001.
[11] See Agresti and Agresti (1978) and Tredoux (1998, 1999) for formulas.

experimental settings (see for example Sauerland & Sporer, 2009; Haw & Fisher, 2004). Overall, the indices summarized here competently and precisely account for parameters of eyewitness lineups. These measures are utilized in the present study.

Lineup Construction Procedures

Steps may be taken to produce a fair lineup from the outset. By considering eyewitness decision-making processes, researchers have proposed general guidelines to lineup construction procedures (Cutler & Kovera, 2010; Luus & Wells, 1991; Turtle et al., 2000; Wells, Rydell & Seelau, 1993). Three such procedures will be summarized here: the *suspect-matched* method, the *perpetrator-description-matched* method, and the *iterative procedure*.

In the suspect-matched method, foils are selected based on their resemblance to the suspect (Cutler & Kovera, 2010). The standard of similarity and acceptability of any particular foil may be a subjective interpretation by a lineup administrator, or a collaborative effort. However, Luus and Wells (1991) note that this procedure can result in a lineup that has foils highly similar to the suspect. This presents an excessively difficult identification task for an eyewitness, even with a clear memory. Perceptual discriminations are necessary for an eyewitness to successfully recognize a suspect from their memory (Wells et al., 1993).The perpetrator-description-matched method permits these discriminations to be made.

The perpetrator-description-matched method involves selecting foils based on the physical description of the suspect provided by an eyewitness. By design, photoarrays constructed using this procedure protect against an eyewitness selecting a suspect from deduction; basing their decision on their own supplied description, as opposed to a clear memory (Cutler & Kovera, 2010). Research has demonstrated that perpetrator-description-matched

photoarrays promote superior rates of accurate identifications to that of suspect-matched photoarrays without increasing false identifications (Wells et al., 1993). However, as noted by Malpass et al. (2007), the fairness of a lineup constructed using this procedure is dependent upon the description from which it is based. If the lineup is constructed based on a vague eyewitness description (e.g., providing only approximations of basic demographic details), the possibility exists to include highly dissimilar foils. The iterative procedure allots discretion to the lineup administrator in the decision to replace and/or remove foils if deemed necessary.

Turtle et al. (2000) proposed an *iterative procedure* to streamline lineup construction techniques. The researchers instructed administrators to: 1) create a desirable pool of lineup foil 'mugshots' from a police database, by selecting photos of individuals that match the witness' description of the suspect; 2) select the foil that is physically closest in appearance to the suspect; 3) place the suspect photograph out of sight; 4) select the second foil of high similarity to the first foil; 5) place the photo of the first foil out of sight; 6) select the third foil as highly similar to the second foil; 7) continue in this manner until you have selected one more foil than required; 8) discard the first foil selected and use remaining foils in the lineup; 9) examine the final lineup, including the suspect, to determine if any foil can be excluded based on the description of the suspect; 10) replace inappropriate foils with an alternative choice, if necessary. The resulting lineup should exhibit *propitious heterogeneity*, or helpful differences, across lineup members (Wells et al., 1993) by containing foils that match the general description of the suspect without possessing an explicit resemblance. The iterative procedure was utilized in constructing the unbiased lineup used in the present study.

Eyewitness Lineups and the Juror

Understanding through *Stinchcombe* that all relevant case materials will appear before legal finders of fact for evaluation, a pertinent question becomes: Are jurors sensitive to foil bias? An opinion questionnaire conducted by Lindsay (1994) illuminated the issue. The questionnaire listed several factors related to eyewitness accuracy and asked participants to rate how likely it is that a witness was accurate under the condition specified, from 1 (*almost certain to be inaccurate*) to 9 (*almost certain to be accurate*). Participants expected eyewitnesses to be significantly less accurate in their decision if presented with an unbiased lineup. Lindsay (1994) noted that evidence suggests eyewitnesses are more likely to be accurate in high-similarity lineups (Lindsay & Wells, 1980; Lindsay et al., 1991). These results suggest that the relationship between lineup fairness and guilt diagnosticity is not intuitively known by jurors. However, questionnaire studies are limited in their ecological application to courtroom proceedings. Simulated trial studies extend the test of layperson knowledge by enrolling participants as jurors, or fact-finders.

Cutler et al. (1988) and Cutler, Penrod and Dexter (1990) manipulated foil bias through fictional testimony.[12] The testimony described the construction and appearance of lineups used in the case and suggested they were either biased or unbiased. Jurors from Cutler et al. (1988) rated the unbiased lineup as significantly fairer than the biased lineup; no such effect was produced in Cutler et al. (1990). In both studies, foil bias produced virtually no effect on jurors' perceived probability the identification was correct. As numerous other independent variables produced

[12]Cutler et al. (1988) recruited undergraduate students as participants in their study. In view of concerns regarding the generalizability of these findings, Cutler et al. (1990) replicated this study, recruiting "eligible and experienced jurors" (p. 186) as participants.

null effects, Cutler et al. (1988) concluded that the results provided "strong evidence that the laypersons do not effectively integrate eyewitness evidence" (p. 54). However, these studies possessed an important limitation by failing to expose jurors to actual photoarray lineups.

Noting that jurors in actual trials are often shown the lineups used as evidence, Devenport, Stinson, Cutler and Kravitz (2002) exposed jurors to a biased or unbiased photoarray lineup. Jurors rated biased lineups as significantly more suggestive and less fair than unbiased lineups, without the aid of expert testimony. Foil bias also produced a significant main effect on defendant culpability, as jurors rated the defendant as more culpable in the unbiased lineup condition than in the biased lineup condition. These results suggest that jurors can recognize bias in lineups and understand that biased lineups increase suggestiveness of the identification procedure. However, these findings have not been robust. Devenport and Cutler (2004) assembled two five-person lineups as biased or unbiased. A hybrid manipulation of these conditions exposed jurors to both descriptive testimony and to the photoarray lineups. Jurors again rated biased lineups as significantly more suggestive than unbiased lineups. Unlike the previous study, foil bias produced no effect on judgments on defendant culpability, nor did it affect perceived accuracy of the eyewitness or verdicts.

Procedures recommended under the *Eyewitness Evidence: A Guide for Law Enforcement* (hereafter the *Guide*) (Technical Working Group, 1999) are considered 'best practice' recommendations for handling eyewitness evidence in the United States. Lampinen, Judges, Odegard and Hamilton (2005) investigated how jurors perceive investigating officers' departures from these practices in a simulated trial. Jurors viewed one of three robbery trial transcripts: 1) no error, wherein the eyewitness simply made an identification from a photo lineup; 2)

procedural error, wherein the defence attorney suggested the investigating officer had made two procedural errors in assembling the lineup (instruction bias and foil bias); and 3) Department of Justice (DOJ) condition, wherein these errors were explicitly challenged by defence counsel as contradictory the *Guide*'s recommendations. Jurors in the DOJ condition rendered the lowest rates of conviction, rendered the lowest credibility ratings of the eyewitness, rated the investigating officer as the least professional, and were more likely to opine that the investigator had jeopardized the prosecution's case. These results suggest that jurors may respond quite negatively to foil bias when presented as a departure from procedure.

In summary, evidence indicates that jurors can independently detect foil bias in eyewitness lineups. However, it is unclear whether lineup type impacts juror decision-making (i.e., verdicts) in minimal trials. Further, few studies have examined whether lineup evidence interacts with PSs and levels of eyewitness confidence. As reviewed, evaluating lineup fairness and following suggested construction procedures allows administrators to draw reliable conclusions regarding the nature of lineup evidence. Study 1 of this two-phased study conducted the mock-witness procedure to provide reliable *a priori* manipulation checks of lineup fairness. Parameters of lineup bias and lineup size were collected for two lineups to support their designations of being biased/unbiased evidence presented to jurors in Study 2.

Study 1: Assessing the Fairness of Eyewitness Lineups

Method

Two eyewitness lineups were constructed and evaluated in a pilot study to validate their use as stimuli in the main study. Study 1 conducted a mock-witness procedure to answer a single research question: Do parameters of the lineup satisfactorily suggest it has met its intended manipulation as biased/unbiased? To answer this question, measurements of *E* and lineup bias were obtained for each lineup. Hypothesis 1 predicted that the evaluation would yield reliable evidence that the lineups were appropriately constructed as unbiased/biased. The iterative procedure was used in the construction of the unbiased lineup.

Participants

Participants ('mock-witnesses') were students enrolled in psychology courses at a university in southwestern Ontario, Canada. The survey-hosting website, Qualtrics[13], documented a total of 47 participants that began the online mock-witness procedure. Data from 2 (4.26%) cases were removed as a result of the survey-hosting website identifying the entries as automated (spam) responses. In addition, a single case (1; 2.13%) was excluded from analysis due to the participant ceasing involvement in the study prior to making a selection from either lineup. The remaining $N = 44$ cases were considered for analysis. On average, mock-witnesses completed the study in approximately 4.5 minutes. Jurors who disclosed recognizing a member of a lineup had their identification data removed only for that particular lineup.

[13] http://www.Qualtrics.com

Out of the 44 mock-witnesses that selected a suspect from the unbiased lineup, 9 (20.0%) self-disclosed recognizing a member of that lineup.[14] These entries were removed. Data from the remaining 35 (n = 35) cases were incorporated into the evaluation. Mock-witnesses were generally of university age (M_{age} = 20.14, *range*: 18 – 45 years, SD = 4.77).[15] Women outnumbered men 30 (85.7%) to 5 (14.3%).[16] The majority identified as White/Caucasian (27; 77.1%). Summary demographic variables are presented in Table 1.

The majority of mock-witnesses that selected from the unbiased lineup also selected from the biased lineup. Data from the biased lineup underwent an identical exclusion process to that of the unbiased lineup. Two (4.55%) of the 44 participants beginning the online mock-witness procedure did not select from the biased lineup. Of the 42 selections from the biased lineup, 9 (21.0%) self-disclosed recognizing a member of the biased lineup. This data was removed. Due to the substantial overlap in mock-witnesses between lineups, demographics from the biased lineup mock-witnesses (n = 33) were highly similar to those obtained from participants in the unbiased lineup (M_{age} = 20.03, *age range*: 18 – 45 years, SD = 4.85).

Materials

Photographs of lineup members were collected from websites hosting public domain 'mugshots' of arrestees. Photographs were collected by defining basic demographic characteristics on a searchable database. In total, 78 photos of *white males, 19-30 years old*, with

[14] Mugshots were collected from public websites.
[15] The vast majority (90.6%) of students enrolled at the University were between the ages of 17 and 25. No enrollment information on race or gender by program was readily available (Gibbons, personal communication, 2014). However, most areas of southwestern Ontario are populated with a White majority (Statistics Canada, 2011).
[16] Women outnumbered men in Study 1 and Study 2. As of January 2014, women outnumbered men 56.6% to 43.4% in the University from which the population was sampled. However, the gender discrepancy from the present study is likely a result of sampling from psychology and criminology classes. As of 2009, there were nearly twice the number of females (124, 074; 11.2% of total enrollment population) to males (66, 918; 6.0%) enrolled in social and behavioural sciences/law university programs in Canada (Statistics Canada, 2010).

Table 1

Mock-Witness Demographics by Lineup Type

	Unbiased Lineup		Biased Lineup	
Demographic	*n*	%	*n*	%
Gender				
Male	5	14.29	5	15.15
Female	30	85.71	28	84.85
Race/ethnicity				
White/Caucasian	27	77.14	24	72.73
Asian/Asiatic-Canadian	2	5.71	2	6.06
Black/African-Canadian	1	2.86	0	0.0
Indian/Indo-Canadian	1	2.86	1	3.03
Latino/Spanish	1	2.86	1	3.03
Other Ethnicity/Combination	1	2.86	2	6.06
Prefer not to disclose	2	5.71	3	9.09
Age				
18-20	26	74.29	25	75.76
21+	9	25.71	8	24.24
Total	35	100	33	100

Note. $N_{UnbiasedLineup} = 35$. $N_{BiasedLineup} = 33$.

brown hair and of *average build* were saved and numbered to functionally serve as the 'police database' selection pool from which the lineups were constructed. The digital photographs were in full-colour and high-quality resolution. For the purposes of the online photoarray presentation, photographs were cropped into squares with dimensions of 180 × 180 pixels, with the face of the lineup member centered in the frame. Lineup members were photographed in front of a grey or grey-blue background and were consistent in overall colouring and presentation.

Two photographs were randomly selected as prototypical suspects from the pool of photos utilizing the SPSS syntax command for random number generation (1-78). The suspect selected for the unbiased lineup was described as a *white male, 19-25 years old, short brown hair, of average build,* with *light facial hair on his chin*. The iterative procedure was followed to build the unbiased lineup. As the degrees of separation grew between the suspect and foil selection, some foils were recognized as differing in physical appearance from the suspect. Specifically, foils were chosen that met all suspect attributes aside from possessing light chin facial hair. Zarkadi, Wade and Stewart (2009) contend that administrators are presented with two options to manage the distinctive feature(s) of a suspect across lineup members: *replication* or *concealment*. The process of replication, outlined below, was conducted with reference to research exploring the *hybrid-similarity* model of recognition (Nosofsky & Zaki, 2003 as cited in Zarkadi et al., 2009).

The hybrid-similarity model assumes that similarity between two faces increases as a "function of the measure of the objects' common features" and decreases as a "function of the measures of the objects' distinctive features" (Knapp, Nosofsky and Busey (2006, p. 878). Across two experiments, Zarkadi et al. (2009) demonstrated that replicating a distinctive feature

across lineup members produces higher rates of correct identifications in target-present lineups than concealment. By replicating distinctive features of a suspect across lineup members, overall familiarity is increased. In turn, this improves the likelihood of an eyewitness recognizing the suspect (Zarkadi et al., 2009). Photographs were digitally manipulated using realistic photo-editing software. To increase the number of effective foils and maintain a meaningful physical description for the mock-witness task, lineup members who did not possess facial hair had facial hair digitally added or darkened in order to effectively match the description of the suspect. In total, 7 of 9 lineup members in the unbiased lineup had facial hair digitally added or darkened.

The suspect in the biased lineup was described as a *white male, 25-30 years old, medium-length brown hair,* and of *average build*. Ineffective foils that differed by at least one primary physical feature were included in the biased lineup. For example, some lineup members failed to meet the described age, hair colour and/or hair length. No lineup members in the biased lineup were digitally manipulated. Lineup members were mutually exclusive to each lineup.

Procedure

The mock-witness procedure was conducted entirely online using the survey-hosting website Qualtrics. Students enrolled in eligible courses under the University Psychology Research Experience Program ('PREP') consented to participate after viewing a study description on a website hosting studies available to be completed for partial course credit. Participants confirmed that they were Canadian citizens at least 18 years of age and therefore legally eligible to serve as jury members. Following consent, participants first viewed an information page. This document outlined that they will read a description of a suspect provided by an eyewitness to a crime, be presented with a set of photographs, and be asked to select the

photograph that they believe is the accused. Participants received additional instructions that they may select only one photograph and that there is no time limit on selection. Participants then viewed a general physical description of a suspect on a single page, as provided above, and were advised that they take as much time as necessary to study the description.

On the following page, participants viewed a simultaneous, nine-person photoarray lineup. Above the lineup, based on instructions provided by Wells and Bradfield (1999), participants were asked "Which person in the lineup do you think is the accused?" Participants made a selection of a single suspect and proceeded to repeat this process for a second lineup. All participants viewed one biased and one unbiased lineup.[17] After viewing and selecting from the second lineup, participants received a debriefing and were thanked for their participation.

Results

Unbiased Lineup

All lineup members in the unbiased lineup received at least 2 (5.71%) selections, suggesting that no null (ineffective) foils appeared in the lineup. The suspect tied with the lineup member in position two for the most selections of any lineup member (7 each; 20.0%). Tredoux' *E* was revealed as 7.34, 95% CI [5.75, 10.13], surpassing the desirable minimum of six recommended by Wells (2001). The nominal size of the lineup (9) fell within the 95% CI, suggesting lineup fairness. Suspect selections did not exceed the 95% CI critical ratio for difference from chance (p (exact) = .052, $z = 1.32$, $p > .05$), suggesting the absence of bias.

[17] Lineup presentation was not counterbalanced and it is possible that identification decisions in the biased lineup were influenced by having already made an identification. There is no data available on the impact of multiple identifications on decision-making (Brewer & Wells, 2006).

Biased Lineup

Two lineup members in the biased lineup were null foils, receiving no selections. The suspect received the majority of mock-witness selections (18; 54.55%). Tredoux' *E* was revealed as 2.90, 95% CI [2.02, 5.19]. The nominal size of the lineup (9) fell outside of the 95% CI, suggesting lineup bias. Further examination revealed substantial bias, as suspect selections were discrepant with values expected from chance (p (exact) = < $.01^{-4}$, $z = 5.01$, $p < .001$). Summary data from the mock-witness procedure is available in Table 2.

Test of Proportions

The McNemar test (McNemar, 1947) is a non-parametric test of the significance of the difference between two correlated groups' proportions.[18] The index tests the null hypothesis of marginal homogeneity: that row totals are equal to corresponding column totals in a 2 × 2 contingency table. To preserve the within-subjects assumption, only decisions from participants that had responded to both lineups without meeting any exclusionary criteria were incorporated into the analysis ($n = 30$). Each decision was recoded into dichotomous outcomes (0 = *failure*, or foil selection; 1 = *success*, or suspect selection). The McNemar test revealed that the proportions of suspect selections significantly differed by lineup type, $p = 0.02$. Overall, 18 (60.0%) participants selected the suspect from the biased lineup, while 7 (23.33%) selected the suspect from the unbiased lineup.

Discussion

Study 1 determined whether each lineup could be reliably demonstrated to be unbiased or biased in order to support their use as stimuli in a minimal trial. Parameters of the unbiased

[18] In small samples, SPSS produces exact results under the cumulative binomial distribution as opposed to the chi-square distribution (Rufibach, 2011 as cited in Adedokun and Burgess, 2012).

Table 2

Frequencies and Percentages of Mock-Witness Selections

Lineup Type	Lineup Member	IDs (%)
Unbiased Lineup	1	2 (5.71)
	2	7 (20)
	3	3 (8.57)
	4	3 (8.57)
	5	3 (8.57)
	6	2 (5.71)
	7	3 (8.57)
	8	5 (14.29)
	9*	7 (20)
Biased Lineup	1	1 (3.03)
	2	2 (6.06)
	3	0 (0)
	4	2 (6.06)
	5	1 (3.03)
	6	0 (0)
	7*	18 (54.55)
	8	5 (15.15)
	9	4 (12.12)

Note. $N_{UnbiasedLineup}$ = 35. $N_{BiasedLineup}$ = 33.
IDs = identifications. * denotes suspect.

lineup support its inclusion as stimuli in Study 2. The provided description of the suspect effectively matched the majority of lineup members. The *E* index was calculated as 7.34, approaching the nominal size of nine and surpassing the desirable minimum of six noted by Wells (2001). The nominal size fell within the 95% CI, suggesting lineup fairness. In addition, the frequency of suspect selections did not exceed critical values of being significantly different from chance, suggesting the absence of bias.

Convincing evidence was found that the biased lineup constructed in this stage is appropriate for use as stimuli in Study 2. As measured through the *E* index, a mere 2.90 lineup members were found to be effective. This finding suggests that at least six members of the lineup (or 66.67%) were determined to be implausible suspects based on the provided description. The majority of mock-witnesses (54.54%) selected the suspect; this observed frequency was extremely unlikely to occur by chance alone. In addition, the McNemar test found that the proportion of suspect selections significantly differed between lineups.

Limitations

The mock-witness procedure is an effective method of collecting identification decisions for assessments of fairness. However, researchers have noted the limits of this paradigm. Corey, Malpass and McQuiston (1999) outline several important considerations. First, mock-witnesses were not provided the option to 'reject' the lineup and the task is, therefore, instruction-biased. Police personnel collecting actual lineup identifications have been advised to provide neutral instructions, indicating that the suspect may not be present in the lineup (Wells et al., 1998). Providing an option to reject the lineup, however, contradicts the goal of the mock-witness task. Collecting decisions from mock-witnesses assists in determining whether a suspect can be

selected based on information available outside that of memory (i.e., the supplied suspect description). How results may have differed if this study incorporated neutral instructions is unknown.

Second, differences exist in the decisions available to mock-witnesses and eyewitnesses. This assessment did not permit mock-witnesses to make multiple selections, which can occur in actual cases. Multiple selections may be treated by an administrator as a suspect selection, foil selection, or a spoiled identification (Corey et al., 1999). Each of these scenarios allow for different conclusions. In addition, the present study, like other mock-witness studies, considered non-choices as "spoiled protocols" (p. S53) and were removed from analysis.

Third, identification decisions made by mock-witnesses are based on entirely different cognitive processes and available information than that of actual eyewitnesses. The written description provided to mock-witnesses is "the most salient information available" (p. S55) to guide a mock-witness choice. Actual eyewitnesses would, of course, attempt to rely on memory in their decision-making. In essence, mock-witnesses are encouraged to make selections based on *relative judgments* by choosing the lineup member who best matches the supplied suspect description, relative to the other lineup members (Wells, 1984). This assists in meeting the underlying goal of the mock-witness procedure of assessing lineup fairness.

Fourth, Lindsay, Smith and Pryke (1999) argue that measures of lineup size and lineup bias are only useful if they postdict identification accuracy. Their findings suggest that lineup bias, but not lineup size, postdict the likelihood of false positive identifications from simultaneous lineups. However, more research is needed into parallelism between mock-witness and eyewitness decisions. As noted by Corey et al. (1999), only when mock-witness

investigations produce outcomes that parallel “stable research findings in the literature using eyewitnesses” (p. S55) can inferences be drawn between mock-witness and eyewitness choosing behaviour.

The results from this present study must be viewed in light of these considerations. However, the concerns of the mock-witness procedure addressed here are primarily in view of the intended function of real-world lineups; i.e., to accurately identify an offender. Study 1 was concerned with reliably demonstrating the status of these lineups as either unbiased or biased. As such, the limitations reviewed do not problematize the methods used in the present study. Overall, convincing evidence was gained that the intended manipulations were successful. The lineups evaluated at this stage were presented as evidence to jurors in a minimal trial.

Study 2: Juror Perceptions of Identification Evidence – An SEO Design

Method

Jurors have been demonstrated to improperly evaluate eyewitness identification evidence. Overall, there is mixed evidence over whether testimony by an expert, or a caution read by a judge, is the most appropriate channel to educate jurors on case-relevant variables. It is imperative to investigate the working Canadian models of both of these procedural safeguards. This effort assists in determining the optimal method to uphold the jury's role as a fact-finder. The present study was undertaken to determine if and how these procedural safeguards influence juror decision-making in cases involving two common and recurring issues raised in cases relying on eyewitness testimony: lineup bias and eyewitness confidence.

Participants

Jurors were recruited from two participant pools. Jurors from the University Psychology Research Experience Program ('PREP') completed the online study for partial course credit (n = 185; 55.06%). The second group of jurors were enrolled as undergraduate Criminology class students ('in-class') at a university in southwestern Ontario and did not receive credit for participation (n = 150; 44.64%). One juror (0.29%) did not identify her participant pool. The trial was completed entirely online using the survey-hosting website Qualtrics.

A total of N = 386 participants were recruited for the study. Only data suitable for analysis was retained. Ten (2.59%) prospective participants ceased involvement in the study immediately following the Case Facts document, indicating some success in excluding participants unable to commit to the role of juror. An additional two (0.52%) prospective participants exited the study prior to the start of the trial, and twenty-nine (7.51%) completed

either a portion or all of the trial, but did not complete any item on the Questionnaire. The overall attrition rate was 10.62%. A further eight (2.07%) prospective participants did not meet age and/or citizenship requirements, and one (0.26%) did not agree to participate in the study. Participants who participated in Study 1 were ineligible to complete Study 2. In total, data from 50 (12.95%) participants were removed, with data from the remaining $N = 336$ participants ("jurors") incorporated into the analysis.

All jurors were required to confirm that they were Canadian citizens at least 18 years of age and therefore legally eligible to serve as jury members. Jurors who disclosed their age ($n = 328$, 97.62%) were generally of university age ($M_{age} = 20.04$, *range*: 18 – 56 years, $SD = 3.56$). Women outnumbered men 236 (70.24%) to 100 (29.76%).[19] The majority identified as White/Caucasian (259; 77.08%). Juror demographics are available in Table 3.

Procedure

After receiving University Research Ethics Board approval, PREP jurors participated in this study by following a link on an online participant recruitment database hosted by the University. In-class jurors completed the online study by following a link posted at the front of class during scheduled lecture time. PREP participants were permitted to begin the study at their own discretion; in-class participants were supervised. A series of independent samples *t*-tests with a rounded Bonferroni correction ($\alpha = .05 / 9 = 0.0055 = \pm 99.45$ CI) revealed that responses on the Questionnaire did not significantly differ by participant type, all *p*s > .05. Data from all jurors was therefore collapsed. The study procedure, outlined below, did not differ between

[19] See above, note 11.

Table 3

Minimal Trial Juror Demographics

Demographic	*n*	%
Gender		
Male	100	29.76
Female	236	70.24
Race/ethnicity		
White/Caucasian	259	77.08
Asian/Asiatic-Canadian	19	5.65
Black/African-Canadian	18	5.36
Indian/Indo-Canadian	12	3.57
Latino/Spanish	6	1.79
Aboriginal/First Nations	3	0.89
Other Ethnicity/Combination	17	5.06
Prefer not to disclose	2	0.60
Age		
18-20	242	72.02
21+	86	25.60
Total	336	100

Note. $N = 336$. $n_{Age} = 328$ (97.62%).

groups. Jurors did not ask any questions or provide feedback before or after the study. On average, jurors completed the study in approximately 17.5 minutes.[20]

Case Facts. Prior to the presentation of evidence, all jurors viewed a succinct case summary entitled Case Facts. The document informed jurors that the accused had pleaded not guilty to the indictable offence of robbery. Jurors were supplied with relevant Canadian Criminal Code provisions under s. 343(3). Jurors read that the case against the defendant rested primarily on a sole eyewitness, who would provide testimony on the circumstances surrounding the robbery. The primary reason charges were laid against the defendant was the positive identification made by the witness from a photographic lineup. Jurors were cautioned to pay careful attention to the testimony and to answer all questions as if they were sitting in an authentic court trial. The document also informed jurors that the trial was based on actual criminal events, that they will be asked to render a verdict following the trial, and that they were permitted to take notes. An exit link was provided for participants that foresaw any reason for being unable to commit to the role of juror.

Eyewitness confidence. Jurors were randomly assigned to one of two eyewitness confidence conditions (low confidence vs. high confidence). Both conditions contained fictional testimony from an assistant manager of a gas station. While fictional, the transcript was constructed with reference to numerous actual in-court examinations of eyewitness. During direct examination, the witness testified that a man entered the store during a summer afternoon, quickly approached the cash register, and demanded money. The witness testified that she stayed

[20] Two jurors' study durations were identified as extreme outliers ($z > 3.29$); these times were excluded from the calculation of the reported average study duration. With these outliers included, the overall average duration increased to approximately 44 minutes. Data from these jurors were retained for the substantive analysis. While the reasons for these prolonged durations are unknown, it is possible that these jurors simply failed to exit the study upon completion.

calm and opened the register. The man then demanded that she “back up” away from the register and threatened to kill her. The witness complied, and raised her hands to demonstrate that she was not activating an alarm. The witness described the offender reaching over the counter and pulling bills out of the register, stuffing them into his pockets, promptly running out of the store and heading left. Viewing conditions inside of the store were described by the eyewitness as “bright”; all lights were working, and it was daytime. Immediately following the robbery, the witness pushed the panic alarm and provided information to a 911 call-taker. No other witnesses were present for the event. The witness explicitly mentioned that no weapon was visible and that she was able to see the man’s face clearly. Security cameras were inactive at the time of the robbery. The witness followed up testimony on the robbery by confirming that the suspect indicated on the lineup viewed by jurors was the person she selected.

Questioning by the Crown attorney led the witness to discuss the photoarray identification. She received a call from an investigating detective and attended the police station. The witness viewed a photographic lineup and made a positive identification of the defendant by pointing to the photo and saying “him”. In the high confidence condition, the witness proclaimed she was “one-hundred percent sure. Positive.” that the defendant committed the robbery. In the low confidence condition, the witness explained she was “pretty sure it was him”.

The defence conducted a brief cross-examination of the witness. The line of questioning was intended to establish doubt of guilt in the mind of jurors (e.g., “Did you ever think that you may have chosen the wrong person? That Mr. Carns is an innocent man?”) When pressed, the witness in the high confidence condition reiterated that she was “one-hundred percent sure it was him” and did “not at all” think she had chosen the wrong person. In the low confidence

condition, the eyewitness' responses were written to indicate a lower level of certainty in the identification in response to the defence's line of questioning. The eyewitness repeated that she was "pretty sure" of her decision and responded with only "no" when asked if she had thought if she had chosen the wrong person. In both conditions, the defence ended questioning by urging jurors to "think hard about this identification".

Lineup type. Jurors viewed the photoarray lineups used in the case immediately following cross-examination of the eyewitness. Jurors were randomly assigned to one of three lineup conditions (no lineup vs. biased lineup vs. unbiased lineup). The nine-person lineups were empirically evaluated for fairness during a pilot stage of this study using the mock-witness procedure. In the biased lineup condition, jurors viewed a lineup that was egregiously biased, in that the provided witness description substantially decreased the effective size of the lineup (Tredoux' E = 2.90, 95% CI [2.02, 5.19]). Jurors in the unbiased lineup condition viewed a lineup composed of members that satisfactorily matched the description provided (Tredoux' E = 7.34, 95% CI [5.75, 10.13)] The McNemar test revealed that proportions of suspect selections significantly differed by lineup type (McNemar p = 0.02). Each simultaneous lineup had a red circle placed around the suspect's photo to indicate the identification made by the witness.

Expert testimony. Jurors were randomly assigned to either view expert testimony or a control (no exposure). The expert testimony transcript was constructed with reference to testimony provided by Elizabeth Loftus in *State v. Goldberg* (2002); the "wig shop murder"; see Lubbock-Avalanche Journal, 2000). Direct examination was conducted by the defence attorney. The expert witness first established her credentials as an expert of eyewitness testimony and memory and provided a general overview of the methods used in eyewitness research. The

substantive portion of the testimony began with outlining the acquisition, retention, and retrieval stages of memory. Next, the expert discussed considerations for proper lineup composition, describing what a fair lineup would look like and the relative judgment decision-making process (Wells, 1984). Finally, the expert discussed the relationship between eyewitness confidence and accuracy, noting that it "can be a very weak relationship" and that it is "common to have people who are very confident and wrong".

Upon cross-examination, the expert acknowledged that some experts in the field of eyewitness testimony disagree about the appropriateness of expert testimony. The prosecution successfully pressed for the expert to acknowledge the common criticism that eyewitness studies possess low ecological validity due to the ethical limitations of research. The expert also acknowledged that an identification made from a biased photoarray lineup may still be correct. Finally, the expert agreed with the formulation put forward by the prosecution that accuracy is more strongly related to confidence for eyewitnesses that make a choice (choosers) as opposed to those that reject a lineup (non-choosers; see Sporer et al., 1995).

Judicial caution. Jurors were randomly assigned to either view the judicial caution or a control (no exposure). Jurors who did not view the caution only viewed standardized instructions pertaining to understanding reasonable doubt and the presumption of innocence. As written by the Canadian Judicial Council (2012), the caution opened by emphasizing that the case depends entirely or largely on identification evidence. The introduction informed jurors that wrongful convictions have occurred due to the reliance placed on mistaken eyewitness identifications. The caution also noted that there is "little connection between great confidence of the witness and the accuracy of the identification". The caution summarized testimony from the eyewitness,

including when the description of the suspect was provided, the "clear and well-lit" viewing conditions, and the close proximity between the offender and the eyewitness. The caution additionally asked jurors to consider how specific the witness' description was of the suspect, how close the description provided was to the way the defendant looked, the fact that the eyewitness had not failed to identify the defendant, and that the eyewitness had not changed her mind about the identification. Several questions about the photoarray lineup were posed to the jury for consideration, on issues surrounding the fairness of the procedure, the suitability of the members selected for the lineup, and the similarity between photos (specifically in size and colour). Immediately following the judicial caution, all jurors viewed the standardized instructions regarding reasonable doubt.

Juror Determination Questionnaire. Following the trial, jurors completed a 10-item questionnaire consisting of single-item assessments. Items for the questionnaire were included with reference to studies that have employed similar evaluations (Cutler et al., 1990; Devenport et al., 2002; Fox & Walters, 1986; Lindsay, Wells & O'Connor, 1989). Jurors first rendered a verdict (guilty or not guilty). Jurors then estimated the percentage of people they believed that would make an accurate identification in the case (0 to 100%). The remaining eight items were all assessed on a 9-point Likert scale. Higher ratings on each item were indicative of greater perceived strength of a construct within that item (e.g., How confident was the eyewitness in this case? 1 = *not at all confident*, 9 = *completely confident*). Jurors rated confidence in their verdict, credibility of the eyewitness, fairness of the overall investigation, how well they understood the trial procedure, the strength of the prosecution's case, and the strength of the defence's case. To serve as manipulation checks, jurors also rated the fairness of the lineup and the confidence of

the eyewitness. Finally, jurors were provided an optional, qualitative content box to submit any notes they had taken during the trial.

Analytical Framework

Design similarities between this study and previous research (Cutler et al., 1989; Devenport et al., 2002) permitted the analysis of data to be undertaken under a similar analytical framework. Juror sensitization is measured through interactions between evidence (eyewitness confidence and lineup type) and the procedural safeguards employed in this study: expert testimony and judicial caution.[21] For example, jurors demonstrate insensitivity to foil bias if no difference exists between group ratings of lineup fairness across lineup type. Jurors who have been sensitized to foil bias by procedural safeguards will rate the fairness of the unbiased lineup significantly higher following exposure to a safeguard. Key dependent variables for assessing juror sensitization are ratings of lineup fairness and ratings of eyewitness credibility.

The key dependent variable for measuring juror skepticism is verdict, an approach employed in previous studies (e.g., Loftus, 1980; Cutler et al., 1990). Juror skepticism is found where a procedural safeguard makes a significant contribution to the logistic regression model to predict verdict. Specifically, jurors will demonstrate skepticism where a safeguard independently and significantly increases not guilty verdicts. Secondary dependent variables for testing of juror skepticism include eyewitness credibility, defence case strength, estimated percentage of accurate identifiers, and prosecution case strength. Evidence for skepticism will be gained where exposure to a safeguard produces significant main effects by decreasing ratings of eyewitness

[21] The study design also allows for testing of an *additive sensitization* effect, wherein sensitivity is induced through exposure to both procedural safeguards.

credibility, decreasing estimates of people who would make an accurate identification, decreasing the strength of the prosecution's case, or increasing the strength of the defence's case.

Juror confusion is evident where jurors exhibit difficulty understanding the scientific evidence presented by the eyewitness expert (Cutler et al., 1989, p. 216). Like sensitization, confusion is measured through interactions, but in directions that contradict the substantive information presented in the content of the testimony or caution. For example, evidence of confusion is gained if jurors exposed to a safeguard rate the credibility of the eyewitness significantly higher in the biased lineup condition compared to jurors who did not view a safeguard. Key dependent variables for measuring juror confusion include ratings of lineup fairness and eyewitness credibility.

Manipulation Checks

A series of *t*-tests with a Bonferroni correction ($\alpha = .05 / 4 = 0.125$) were conducted to ensure that manipulations were successful and that jurors understood the trial procedure. Satterthwaite's (1946) approximation for degrees of freedom was used where equal variances were not assumed.[22] Cohen's *d* is reported as an effect size for independent samples *t*-tests, with values of 0.2, 0.5 and 0.8 indicating small, medium, and as large effect sizes, respectively (Cohen, 1988).[23]

Eyewitness lineups. Jurors were either exposed (214, 63.69%) or not exposed (122, 36.31% to an eyewitness lineup. In total, 114 (33.93%) viewed an unbiased lineup and 100

[22] A *t-test* is highly robust under violations of normality when sample sizes are large (Boneau, 1960; Sawilowsky & Blair, 1992). However, violations of the assumption of homogeneity of variance can yield unreliable test statistics (Levene, 1960; Ramsey, 1980). Satterthwaite's correction uses the chi-square distribution to approximate the degrees of freedom from two independent sample variances where equal variances cannot be assumed.

[23] The *d* index is calculated as $d = (M_1 - M_2) / s_{Pooled}$, where M_1 is the mean of the first group, M_2 is the mean of the second group, and s_{Pooled} is the pooled standard deviation of the data.

(29.76%) viewed a biased lineup. Jurors rated the fairness of the lineup from 1 (*not at all fair*) to 9 (*completely fair*). Levene's test revealed that the assumption of homogeneity of variance was violated, $F(2, 212) = 5.23$, $p < .05$. Therefore, equal variances were not assumed. Jurors rated the unbiased lineup ($M = 6.25$, $SD = 1.88$) fairer than the biased lineup ($M = 5.15$, $SD = 2.28$), supporting the manipulation. This difference was significant, $t(192.25) = 3.81$, $p < .001$ and represented a medium effect size ($d = 0.55$).

Eyewitness confidence. Eyewitness testimony was presented in a condition of either low confidence ($n = 163$, 48.51%) or high confidence ($n = 173$, 51.49%). Jurors rated the confidence of the eyewitness from 1 (*not at all confident*) to 9 (*completely confident*). Levene's test revealed that the assumption of homogeneity of variance was violated, $F(2, 334) = 4.84$, $p < .05$. Therefore, equal variances were not assumed. Supporting the manipulation, jurors rated the eyewitness expressing high confidence ($M = 7.91$, $SD = 1.61$) as more confident than the eyewitness expressing low confidence ($M = 6.08$, $SD = 1.90$). This difference was significant, $t(318.31) = 9.51$, $p < .001$ and represented a large effect size ($d = 1.07$).

Procedural safeguards. Jurors were either exposed or not exposed to procedural safeguards of expert testimony and/or judicial caution. All jurors exposed to a procedural safeguard (judicial caution $n = 159$, 47.32%; expert testimony $n = 153$, 45.54%) indicated the transcript had been viewed by selecting a labelled radial button presented following the text. Jurors reported how well they understood the trial procedure from 1 (*did not at all understand*) to 9 (*completely understood*). Overall, participants reported a high perceived understanding of the trial procedure ($M = 7.46$, $SD = 1.65$). A Kruskal-Wallis test indicated that understanding of the trial between jurors who viewed expert testimony ($M = 7.31$, $SD = 1.65$) and who did not view

expert testimony (M = 7.59, SD = 1.64) approach significance, $H(1) = 3.80$, $p = .05$, $d = 0.21$. Understanding of the trial procedure did not differ between jurors who did not view the judicial caution (M = 7.38, SD = 1.73) and who viewed the judicial caution (M = 7.56, SD = 1.56, $H(1) = 0.59$, $p > .05$, $d = .08$).

Summary

Manipulation checks conducted on the independent variables in this study indicated that eyewitness confidence, lineup type, expert testimony, and judicial caution were successful. In addition, jurors' ratings of perceived comprehension of the trial procedure were high across all conditions. All independent variables were retained for the analyses to follow.

Hypotheses

Hypotheses were constructed based on the three plausible effects on jurors from evidence presented in this design: skepticism, sensitivity, and confusion. Based on literature reviews by Martire (2008) and Martire and Kemp (2011), hypothesis 1 predicted that expert testimony would produce juror skepticism. Expert testimony was expected to significantly predict verdict in a logistic regression model. Specifically, jurors exposed to expert testimony were anticipated to be more likely to render not guilty verdicts.

Based on results from Devenport et al. (2002), hypothesis 2 predicted that lineup type would produce a significant main effect on jurors' ratings of investigation fairness. Specifically, jurors who viewed an unbiased lineup were anticipated to render significantly higher ratings of investigation fairness than jurors exposed to the biased lineup. Hypothesis 3 anticipated juror sensitization through a significant lineup type × expert testimony interaction on the credibility of the eyewitness. Specifically, jurors exposed to expert testimony who viewed a biased lineup

were expected to estimate a significantly lower rating of eyewitness credibility compared to jurors who viewed the unbiased lineup.

Based on results from Wells et al. (1979) and Whitley and Greenberg (1986), hypothesis 4 predicted that eyewitness confidence would have a significant main effect on jurors' perceived credibility of the eyewitness. Specifically, jurors exposed to testimony in the high confidence condition would rate the credibility of the eyewitness significantly higher than jurors from the low confidence condition. Finally, hypothesis 5 predicted that the judicial caution would render null main effects on all questionnaire items and fail to sensitize jurors through interactions.

Results

Verdict

Following the trial, each juror independently rendered a verdict in the case. To ascertain whether any independent variable contributed to predicting jurors' verdicts, a binary logistic regression was conducted with verdict as the dichotomous dependent variable (0 = *not guilty*, 1 = *guilty*). All predictors were included in the logistic regression model and dummy-coded to assign control categories (e.g., 0 = *no expert testimony*, 1 = *expert testimony*).[24] Due to previous research identifying eyewitness confidence, lineup type, expert testimony and judicial cautions as impactful on jurors' appraisals of defendant culpability and/or verdicts, all were included in a forced entry logistic regression model.

Model diagnostic statistics were interpreted following guidelines provided by Field (2009) to assess points of poor fit and undue influence. Cook's Distances (Cook, 1977) were all less than 1 (*range* = .01 – .05, *M* = .02), indicating no points of undue influence (according to

[24] Data from jurors who were not exposed to a lineup was removed from this analysis because it was of no theoretical interest. Lineups were dummy-coded (0 = *biased lineup*, 1 = *unbiased lineup*) and included in the model.

Cook and Weisberg, 1982). DFBeta values for the constant (*range* = -.05 – .04) and for all predictors (*range* = -.04 – .04) were also under the threshold of 1. Individual cases fell very close to their expected leverage value, being $k + 1 / N = 5 / 214 = .02$ (*range* = .02 – .03, *M* = .02). Overall, no cases were isolated that possessed undue influence on the model. No points of poor fit were identified, as all normalized residuals fell outside of the critical value of ±1.96 (*range* = -1.46 – 1.53).

In total, $n = 214$ verdicts were included in the model. Verdicts were virtually split overall; 109 jurors (50.93%) rendered verdicts of not guilty and 105 jurors (49.07%) rendered verdicts of guilty. The baseline model, therefore, estimated that all jurors rendered verdicts of not guilty (50.93% correct classification, *-2LL* = 296.59). Including all variables in the model increased the percentage of correctly classified responses to 58.88%. Variable coefficients were determined to significantly increase the predictive power of the model overall (*-2LL* = 285.96, model $\chi^2 = 10.64$, $p < .05$). To assess overall goodness-of-fit (GOF), Nagelkerke's (1991) R_N^2 was calculated as .07, Cox and Snell's (1989) R^2 was calculated as .05, and Hosmer and Lemeshow's (1989) R_L^2 was calculated as .04. GOF values indicated that including the predictors produced minimal improvement in the ability of the LR model to predict verdict, compared to that of the intercept. Wald statistics were calculated to evaluate individual predictors. When this index yields statistical significance for a particular variable, evidence is gained that the predictive power of the model is increased with the inclusion of that variable (Crichton, 2001). This analysis revealed that eyewitness confidence (Wald χ^2 (1) = 0.45, $p > .05$), judicial caution (Wald χ^2 (1) = 2.37, $p > .05$ and, contrary to hypothesis 1, expert testimony (Wald χ^2 (1) = 0.54, $p > .05$) did not make significant contributions to the predictive power of the model. Lineup type, however, did

significantly contribute to the ability of the model to predict verdict (Wald χ^2 (1) = 7.48, $p < .01$, *OR* = .46, 95% CI [.27, .80]; see Table 4). Overall, 61% of jurors who viewed a biased lineup rendered a not guilty verdict and 39% rendered a guilty verdict. Of jurors who viewed an unbiased lineup, 42.11% rendered a not guilty verdict and 57.89% rendered a guilty verdict.

Questionnaire Hypothesis Testing

A series of analyses of variance (ANOVA) were conducted on jurors' responses from the Juror Determination Questionnaire. Unless otherwise noted, all tests conducted were four-way independent ANOVAs with a 3 (lineup exposure: no lineup vs. unbiased lineup vs. unbiased lineup) × 2 (eyewitness confidence: low vs. high) × 2 (expert testimony: no exposure vs. exposure) × 2 (judicial caution: no exposure vs. exposure) factorial design. Alpha was set at $p < .05$ for all analyses. Partial eta-squared (η_p^2) is reported as an effect size measure.[25] Summary descriptive statistics for each questionnaire item are reported in Table 5.

Investigation fairness. Jurors rated the overall fairness of the investigation on a scale from 1 (*not at all fair*) to 9 (*completely fair*). Levene's test revealed that the assumption of homogeneity of variance was tenable, $F(24, 312) = 1.32$, $p > .05$. Supporting hypothesis 2, there was a significant main effect of lineup type on ratings of investigation fairness, $F(2, 334) = 6.10$, $p < .01$, $\eta_p^2 = .04$, $\pi = .89$. Jurors exposed to an unbiased lineup rated the overall fairness of the

[25]The effect size η_p^2 is calculated as $SS_{between} / SS_{between} - SS_{error}$ (Levine & Hullett, 2002). This measure accounts for the proportion of unique variance of an independent variable on a dependent variable, producing a value "from which all other nonerror sources of variance in the experiment have been removed" (Cohen, 1973, p. 108). Partial eta squared is non-additive and susceptible to being an upwardly biased estimate, particularly with small sample sizes (Pierce, Block & Aguinis, 2004). Therefore, reported values must be interpreted with caution when following guidelines recommended by Cohen (1988) of .01, .06, and .14 as indicative of small, medium and large effect sizes, respectively.

Table 4

Results of Logistic Regression on Verdict

	β (SE)	Lower	Odds Ratio	Upper
Constant	.11 (0.30)			
Eyewitness Confidence	-.19 (0.28)	0.46	0.83	1.44
Lineup Type	-.77* (0.28)	0.27	0.46	0.80
Expert Testimony	.21 (0.28)	0.71	1.23	2.15
Judicial Caution	.44 (0.29)	0.89	1.55	2.71

Note. R_N^2 = .07 (Nagelkerke). R^2 = .05 (Cox & Snell), R_L^2 = .04 (Hosmer & Lemeshow). Model χ^2 = 10.64, $p < .05$. * $p < .01$.

Table 5

Descriptive Statistics on Questionnaire Items by Independent Variable

	Juror Determination Questionnaire Item									
	What is your verdict in this case?[1]	How confident are you in your verdict?[2]	How credible did you find the eyewitness in the case?[3]	How confident was the eyewitness in this case?[4]	How fair was the lineup in this case?[5]	How well did you understand the trial procedure?[6]	Overall, did the investigating officers in this case conduct a fair investigation?[7]	How strong is the prosecution's case?[8]	How strong is the defense's case?[9]	What percentage of people do you think would make an accurate identification in this case?[10]
Condition (*n*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
EC										
LC (163)	.50 (.50)	6.34 (1.71)	5.91 (2.02)	6.08 (1.90)	5.67 (2.10)	7.47 (1.56)	6.17 (1.62)	5.90 (1.74)	5.58 (1.63)	58.37 (20.17)
HC (173)	.49 (.50)	6.27 (1.62)	5.89 (1.93)	7.91 (1.61)	5.72 (1.96)	7.46 (1.73)	6.02 (1.92)	5.74 (1.96)	5.23 (1.90)	55.94 (20.72)
ET										
NET (183)	.50 (.50)	6.36 (1.60)	5.91 (1.97)	7.19 (1.88)	5.57 (2.15)	7.59 (1.64)	5.95 (1.88)	5.72 (1.95)	5.01 (1.82)	57.11 (21.55)
ET (153)	.50 (.50)	6.25 (1.73)	5.89 (1.97)	6.82 (2.07)	5.84 (1.87)	7.31 (1.65)	6.27 (1.64)	5.93 (1.73)	5.86 (1.62)	57.13 (19.15)
LT										
NL (122)	.51 (.50)	6.17 (1.82)	5.79 (2.01)	7.02 (2.08)	5.63 (1.82)	7.20 (1.76)	6.02 (1.80)	5.60 (1.81)	5.55 (1.76)	58.88 (19.96)
BL (100)	.39 (.49)	6.39 (1.63)	5.77 (1.84)	6.89 (1.87)	5.15 (2.28)	7.60 (1.54)	5.67 (1.96)	5.85 (1.86)	5.18 (1.77)	55.38 (20.99)
UL (114)	.58 (.50)	6.38 (1.51)	6.13 (2.03)	7.14 (1.96)	6.25 (1.88)	7.62 (1.55)	6.54 (1.48)	6.03 (1.89)	5.42 (1.90)	56.75 (20.56)
JC										
NJC (177)	.48 (.50)	6.25 (1.75)	5.76 (2.07)	6.84 (1.98)	5.76 (2.15)	7.38 (1.76)	5.95 (1.84)	5.74 (1.81)	5.46 (1.78)	57.44 (21.54)
JC (159)	.52 (.50)	6.36 (1.57)	6.05 (1.85)	7.23 (1.96)	5.62 (1.88)	7.56 (1.56)	6.25 (1.70)	5.91 (1.90)	5.33 (1.78)	56.71 (19.25)

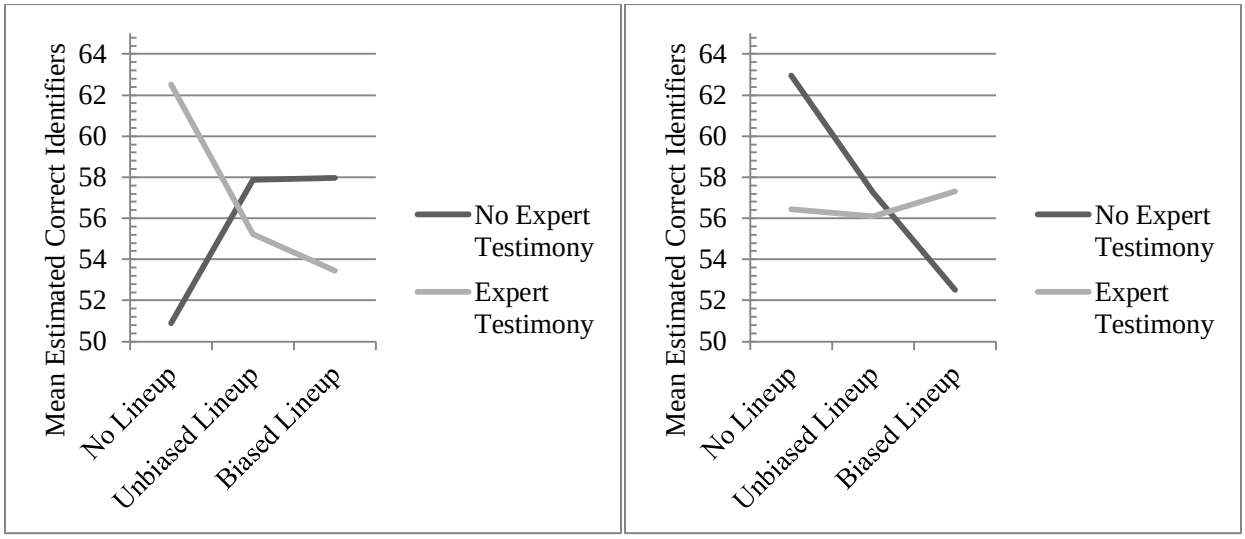
Note. *N* = 336. 1 = (0 = *Not guilty*, 1 = *Guilty)*, 2 = (1 = *Not at all confident*, 9 = *Completely confident*) 3 = (1 = *Not at all credible*, 9 = *Completely credible*), 4 = (1 = *Not at all confident*, 9 = *Completely confident)*, 5 = (1 = *Not at all fair*, 9 = *Completely fair*), 6 = (1 = *Did not at all understand*, 9 = *Completely understood*), 7 = (1 = *Not at all fair*, 9 = *Completely fair*), 8 = (1 = *Not at all strong*, 9 = *Completely strong*), 9 = (1 = *Not at all strong*, 9 = *Completely strong*), 10 = (0 - 100%), EC = *Eyewitness confidence*, LC = *Low confidence*, HC = *High confidence*, ET = *Expert testimony*, NET = *No expert testimony*, LT = *Lineup type*, NL = *No lineup*, BL = *Biased lineup*, UL = *Unbiased lineup*, JC = *Judicial caution*, NJC = *No judicial caution.*

investigation significantly higher (*M* = 6.57, *SD* = .17) than jurors who did not view a lineup (*M* = 6.08, *SD* = .17 *p* < .05) and jurors who viewed a biased lineup (*M* = 5.72, *SD* = .18, *p* < .001).There was also a significant eyewitness confidence × lineup type interaction, *F*(2, 334) = 3.31, *p* < .05, η_p^2 = .02, π = .63. In the high confidence condition, jurors who viewed an unbiased lineup rated the overall fairness of the investigation significantly higher (*M* = 6.84, *SD* = .24) than jurors who did not view a lineup (*M* = 5.74, *SD* = .21, *p* < .001) and jurors who viewed a biased lineup (*M* = 5.63, *SD* = .26, *p* < .001). Ratings of investigation fairness from jurors in the low confidence condition did not significantly differ by lineup type (no lineup *M* = 5.60, *SD* = .24; unbiased lineup *M* = 6.42, *SD* = .28; biased lineup *M* = 5.16, *SD* = .30, *p*s > .05).

Estimated percentage of accurate identifiers. Jurors estimated the percentage of people who would make an accurate identification in the case on a scale from 0 to 100%. Levene's test revealed that the assumption of homogeneity of variance was tenable, *F*(24, 312) = 1.44, *p* > .05. There was a significant lineup type × expert testimony × judicial caution interaction, *F*(2, 334) = 3.49, *p* < .05, η_p^2 = .02, π = .65. Follow-up simple effects analysis revealed two significant three-way interactions. First, in the absence of both procedural safeguards, jurors who did not view a lineup estimated a significantly higher percentage of people would make an accurate identification in the case (*M* = 62.97, *SD* = 3.28) than jurors who viewed a biased lineup (*M* = 52.52, *SD* = 3.87, *p* < .05). Second, jurors who viewed a judicial caution absent a lineup rendered significantly lower estimates when not exposed to expert testimony (*M* = 50.89, *SD* = 4.10) compared to jurors exposed to expert testimony (*M* = 62.52, *SD* = 3.74, *p* < .05). Results for the three-way ANOVAs are presented in Figure 1.

Figure 1

Three-way ANOVA on Estimated Percentage of Correct Identifiers by Judicial Caution, Lineup Type, and Expert Testimony



Judicial Caution No Judicial Caution

Note. N = 336. What percentage of people do you think would make an accurate identification in this case? (0 - 100%).

Verdict confidence. Jurors rated the confidence in their verdict on a scale from 1 (*not at all confident*) to 9 (*completely confident*). As a preliminary analysis, an independent samples *t*-test revealed that jurors who rendered a guilty verdict were significantly more confident in their verdicts (M = 6.72, SD = 1.44) than jurors who rendered a not guilty verdict (M = 5.90, SD = 1.77), $t(2, 321.64) = 4.66$, $p < .001$, $d = 0.52$. Four-way independent factorial ANOVAs were conducted on groups isolated by verdict (not guilty n = 169, 50.30%; guilty n = 167, 49.70%). Results are reported for jurors who rendered a guilty verdict (n = 167).[26]

Levene's test revealed that the assumption of homogeneity of variance was tenable, $F(24, 143) = 1.46$, $p > .05$. There was a significant main effect of eyewitness confidence on verdict confidence, $F(1, 166) = 7.17$, $p < .01$, $\eta_p^2 = .05$, $\pi = .76$. Interestingly, pairwise comparisons revealed that jurors were marginally more confident in their guilty verdict having viewed an eyewitness with low confidence (M = 6.94, SD = .16) than an eyewitness with high confidence (M = 6.51, SD = .16, p = 0.51). However, this effect was qualified by a significant lineup type × eyewitness confidence interaction, $F(2, 165) = 4.72$, $p < .05$, $\eta_p^2 = .06$, $\pi = .78$. Jurors in the high confidence condition who viewed an unbiased lineup (M = 7.03, SD = .26) were significantly more confident in their guilty verdicts than jurors who viewed a biased lineup (M = 6.06, SD = .33, $p < .05$) and jurors who did not view a lineup (M = 6.30, SD = .23, $p < .05$). Guilty verdict confidence did not significantly differ for jurors exposed to low confidence across all lineup conditions (no lineup M = 7.08, SD = .28; unbiased lineup M = 6.69, SD = .23; biased lineup M = 7.19, SD = .31, $p > .05$).

[26] There were no significant main effects or interactions on verdict confidence for jurors who rendered a not guilty verdict, all ps > .05.

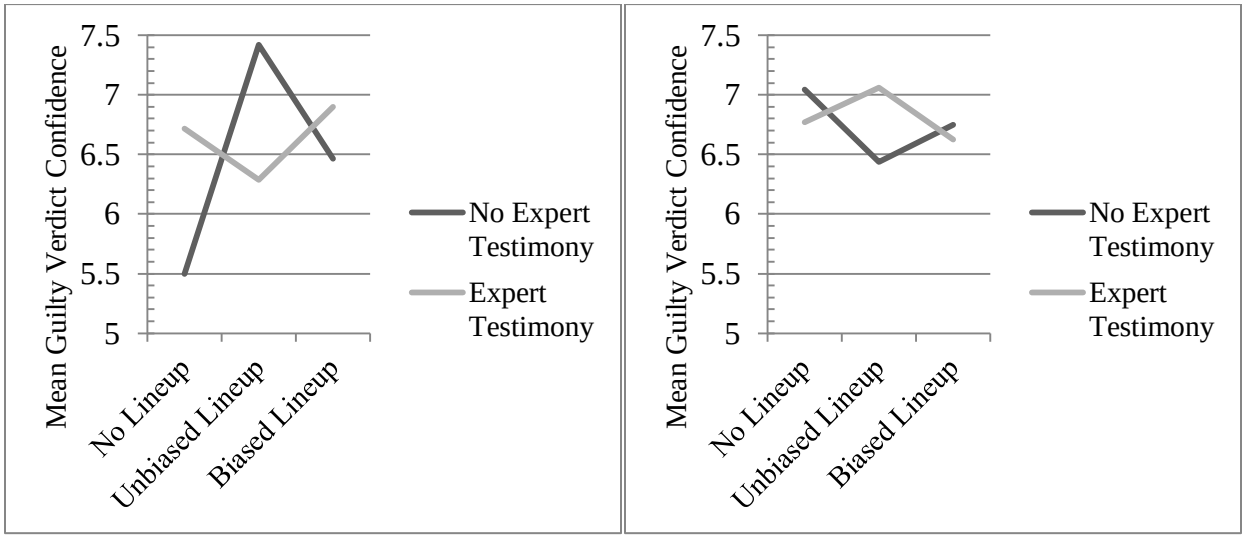
There were also significant lineup type × expert testimony × judicial caution interactions, $F(2, 165) = 402$, $p < .01$, $\eta_p^2 = .05$, $\pi = .71$. Jurors who did not view a lineup or expert testimony were significantly more confident in their guilty verdicts without viewing a judicial caution ($M = 7.04$, $SD = .29$) than with the judicial caution ($M = 5.50$, $SD = .41$, $p < .01$). In addition, jurors who viewed the unbiased lineup without expert testimony were significantly more confident in their guilty verdicts with the judicial caution ($M = 7.42$, $SD = .32$) than without the judicial caution ($M = 6.44$, $SD = .35$, $p < .05$). Results are presented in Figure 2.

Eyewitness credibility. Jurors rated the credibility of the eyewitness on a scale from 1 (*not at all credible*) to 9 (*completely credible*). Levene's test revealed that the assumption of homogeneity of variance was tenable, $F(24, 312) = 1.44$, $p > .05$. Contrary to hypothesis 3, there was no significant lineup type × expert testimony interaction, though this test yielded inadequate statistical power, $F(2, 334) = .21$, $p > .05$, $\eta_p^2 = .00$, $\pi = .08$. *Post hoc* tests supported that jurors' ratings of eyewitness credibility did not significantly differ across lineup type, for jurors exposed to expert testimony (no lineup $M = 5.86$, $SD = .26$; unbiased lineup $M = 6.00$, $SD = .27$; biased lineup $M = 5.79$, $SD = .30$) or not exposed to expert testimony (no lineup $M = 5.72$, $SD = .25$; unbiased lineup $M = 6.24$, $SD = .25$; biased lineup $M = 5.75$, $SD = .26$).

Contrary to hypothesis 4, eyewitness confidence did not produce a significant main effect on ratings of eyewitness credibility, $F(1, 335) = .09$, $p > .05$, $\eta_p^2 = .00$, $\pi = .06$ (low confidence $M = 5.91$, $SD = 2.02$; high confidence $M = 5.89$, $SD = 1.93$). However, simple effects analysis revealed a significant eyewitness confidence × lineup type interaction. In the high confidence condition, jurors rated the credibility of the eyewitness significantly higher when having selected the suspect from an unbiased biased lineup ($M = 6.22$, $SD = .28$) compared to a biased lineup,

Figure 2

Three-way ANOVA on Guilty Verdict Confidence by Lineup Type, Judicial Caution, and Expert Testimony



Judicial Caution No Judicial Caution

Note. n = 167. How confident are you in your verdict? (1 = *Not at all confident*, 9 = *Completely confident*).

(*M* = 5.49, *SD* = .29 $p < .05$; see Figure 3).

Lineup fairness. Jurors rated the fairness of the lineups on a scale from 1 (*not at all fair*) to 9 (*completely fair*). Levene's test revealed that the assumption of homogeneity of variance was violated, $F(24, 312) = 2.21$, $p < .01$.[27] As a result, data from jurors who did not view a lineup was removed and juror sensitivity was tested by using two 2 × 2 independent factorial ANOVAs.

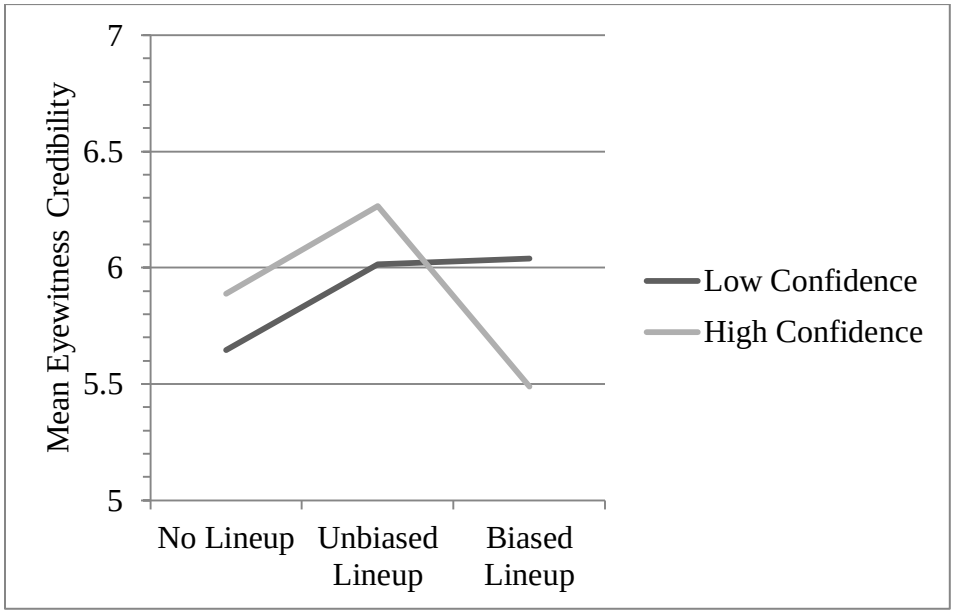
To test for lineup fairness ratings across judicial caution and lineup type, a 2 (judicial caution: no exposure vs. exposure) by 2 (lineup type: unbiased lineup vs. biased lineup) independent factorial ANOVA was conducted assuming equal variances, $F(4, 210) = 2.08$, $p > .05$. As expected, lineup type produced a significant main effect on jurors' ratings of lineup fairness, $F(1, 213) = 14.59$, $p < .01$, $\eta_p^2 = .07$, $\pi = .97$. The interaction term provided no evidence that the judicial caution sensitized jurors to the fairness of the lineups, $F(1, 213) = 2.97$, $p > .05$, $\eta_p^2 = .01$, $\pi = .40$. Ancillary analyses were conducted due to the low statistical power derived for this specific test.

Partial evidence was found that the judicial caution moderated the zero-order correlation between lineup type and lineup fairness ratings (see Baron & Kenny, 1986). A point-biserial correlation demonstrated that jurors' ratings of lineup fairness were significantly associated with exposure to an unbiased lineup compared to a biased lineup, $r_{pb} = .26$, $p < .001$. This significant positive association existed for jurors who did not view a judicial caution, $r_{pb} = .35$, $p < .001$, but not for jurors exposed to a judicial caution, $r_{pb} = .15$, $p > .05$, and the difference between these two groups' correlation coefficients approached significance, $z = 1.53$, p (one-tailed) = .06.

[27]This assumption was also violated using 3-way ANOVAs.

Figure 3

Two-way ANOVA on Ratings of Eyewitness Credibility by Lineup Type and Eyewitness Confidence



Note. *N* = 336. How credible did you find the eyewitness in the case? (1 = *Not at all credible*, 9 = *Completely credible*).

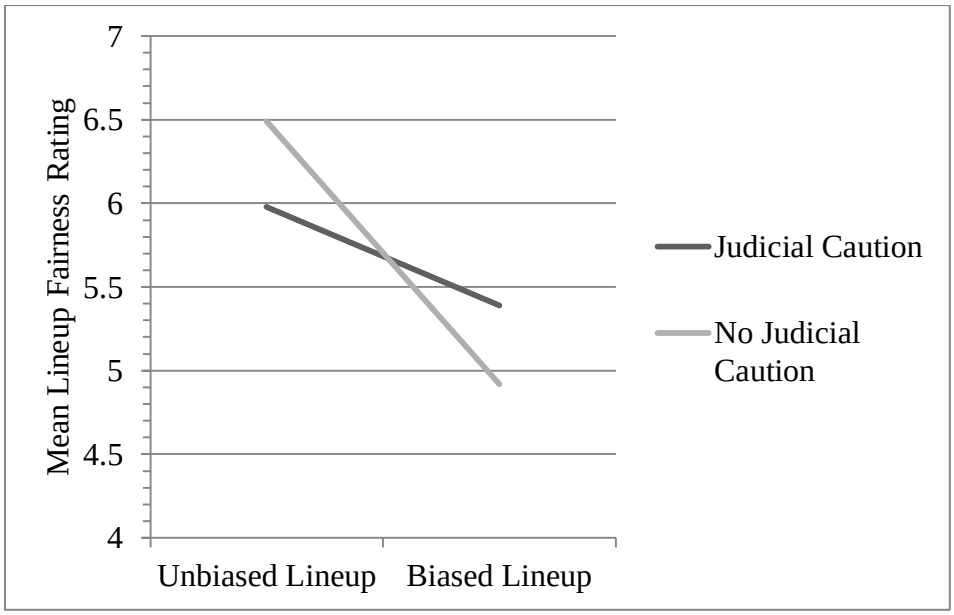
Further, independent samples *t*-tests revealed that jurors who did not view the judicial caution correctly appraised the unbiased lineup as significantly more fair (M = 6.49, SD = 1.93) than the biased lineup (M = 4.92, SD = 2.35, $p < .001$). However, lineup fairness ratings from jurors who viewed the judicial caution did not differ between the unbiased lineup (M = 5.98, SD = 1.79) and biased lineup conditions (M = 5.39, SD =2.20, $p > .05$; see Figure 4).

To test for lineup fairness ratings across expert testimony and lineup type, a 2 (expert testimony: no exposure vs. exposure) by 2 (lineup type: unbiased lineup vs. biased lineup) independent factorial ANOVA was conducted assuming equal variances, $F(4, 210) = 2.40$, $p > .05$. The interaction term and *post hoc* tests provided no evidence that expert testimony impacted jurors' ratings of lineup fairness across lineup type, $F(1, 213) = .34$, $p > .05$, $\eta_p^2 = .00$, $\pi = .09$ (no expert testimony: unbiased lineup M = 6.15, SD = .25, biased lineup M = 4.91, SD = .26; expert testimony: unbiased lineup M = 6.37, SD = .28; biased lineup M = 5.47, SD = .30). An ancillary correlational analysis was conducted due to the low statistical power for this particular test, but did not reveal evidence of moderation effects.

Defence case strength. Jurors rated the strength of the defence's case from 1 (*not at all strong*) to 9 (*completely strong*). Levene's test revealed that the assumption of homogeneity of variance was tenable, $F(24, 312) = 1.50$, $p > .05$. Expert testimony produced a significant main effect on jurors' ratings of the strength of the defence's case, $F(1, 335) = 17.88$, $p < .001$, $\eta_p^2 = .05$, $\pi = .99$. Jurors exposed to expert testimony rated the strength of the defence's case significantly higher (M = 5.83, SD = .14) than jurors who did not view expert testimony (M = 4.99, SD = .14). There were no other significant effects or interactions.

Figure 4

Two-way ANOVA on Ratings of Lineup Fairness by Lineup Type and Judicial Caution



Note. *n* = 214. How fair was the lineup in this case? (1 = *Not at all fair*, 9 = *Completely fair*).

Prosecution case strength. Jurors rated the strength of the prosecution's case from 1 (*not at all strong*) to 9 (*completely strong*). Levene's test revealed that the assumption of homogeneity of variance was violated, $F(24, 312) = 1.70$, $p < .05$. As a result, two 3 × 2 × 2 independent factorial ANOVAs were conducted.

The effects of the judicial caution was tested by conducting a 3 (lineup type: no lineup vs. biased lineup vs. unbiased lineup) × 2 (judicial caution: no exposure vs. exposure) × 2 (eyewitness confidence: low confidence vs. high confidence) ANOVA assuming equal variances, $F(12, 324) = 1.79$, $p > .05$. The omnibus test revealed no significant main effects or interactions. However, *post hoc* tests revealed that, in the high confidence condition and absent a judicial caution, jurors rated the strength of the prosecution's case significantly higher having viewed an unbiased lineup ($M = 6.21$, $SD = .38$) relative to biased lineup ($M = 5.09$, $SD = .40$, $p < .05$). Three-way ANOVA results are presented in Figure 5.

The effects of expert testimony was tested by conducting a second 3 (lineup type: no lineup vs. biased lineup vs. unbiased lineup) × 2 (expert testimony: no exposure vs. exposure) × 2 (eyewitness confidence: low confidence vs. high confidence) ANOVA, also finding the assumption of equal variances tenable, $F(12, 324) = .78$, $p > .05$. There were no significant main effects or interactions. Gabriel's pairwise *post hoc* tests revealed no significant effects. Three-way ANOVA results are presented in Figure 6.

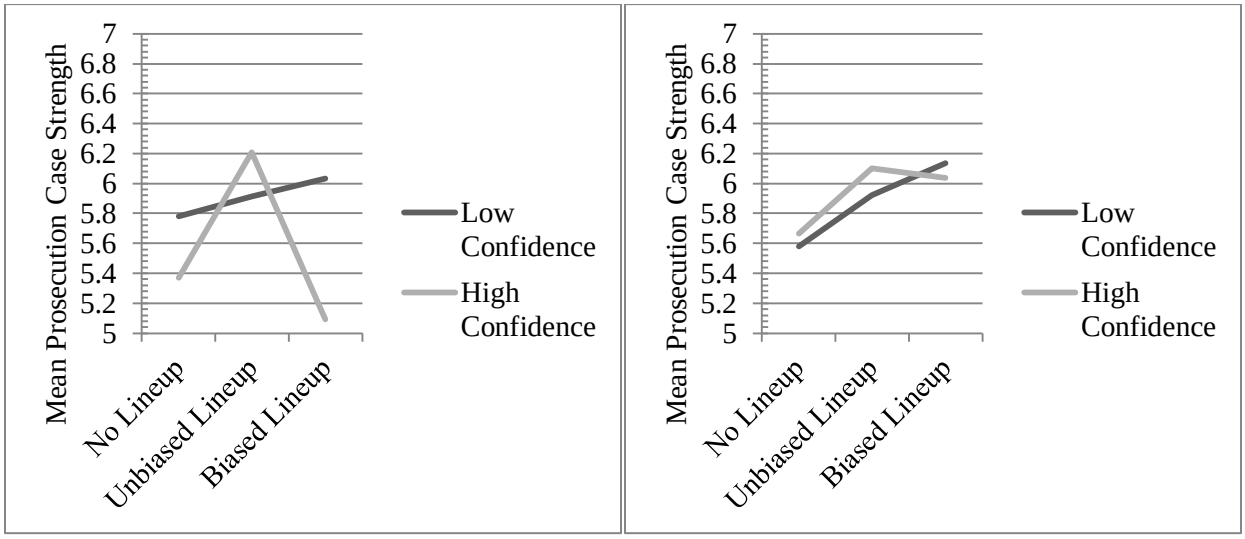
Discussion

Assessing the Influence of Eyewitness Confidence

Results from the present study contradict results from numerous studies that have identified eyewitness confidence as highly influential on jurors' judgments of a defendant's guilt.

Figure 5

Three-way ANOVA on Prosecution Case Strength by Lineup Type, Judicial Caution, and Eyewitness Confidence

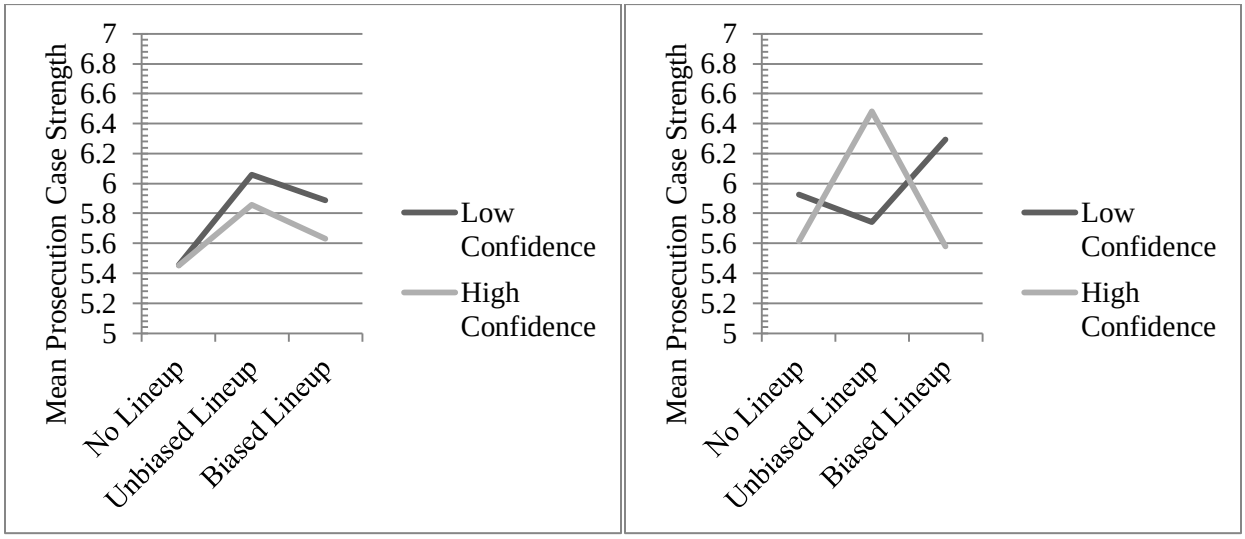


No Judicial Caution Judicial Caution

Note. *N* = 336. How strong is the prosecution's case? (1 = *Not at all strong*, 9 = *Completely strong*).

Figure 6

Three-way ANOVA on Prosecution Case Strength Ratings by Lineup Type, Expert Testimony, and Eyewitness Confidence



No Expert Testimony Expert Testimony

Note. N = 336. How strong is the prosecution's case? (1 = *Not at all strong*, 9 = *Completely strong*).

Eyewitness confidence did not significantly predict verdict in a logistic regression model. The baseline model correctly classified 50.9% of verdict estimates, which was approximated by jurors regardless of whether they viewed an eyewitness with low confidence (49.7% not guilty to 50.3% guilty) or high confidence (50.9% not guilty to 49.1% guilty). Further, eyewitness confidence also did not directly impact jurors' ratings of eyewitness credibility. Overall, jurors believed that the eyewitness testimony was only moderately credible; a trend that continued in conditions of low confidence and high confidence. On the grounds that the manipulation of confidence was successful, these null effects cumulatively indicate that jurors did not afford undue probative weight to the eyewitness testimony.

Higher-order interactions indicated that jurors considered other evidence in combination with confidence in forming judgments. Specifically, findings from the present study indicate that jurors: a) rated the credibility of a highly confident witness higher when that witness had selected the defendant from an unbiased lineup, rather than a biased lineup; b) were more confident in their guilty verdicts when a highly confident witness had selected from an unbiased lineup, rather than a biased lineup; and c) rated investigation fairness higher when a highly confident witness selected from an unbiased lineup, rather than no lineup or a biased lineup. These results provide promising indications of jurors' abilities to integrate diverse evidence into decision-making. It is clear that high eyewitness confidence still had some persuasive effect on jurors; however, this was only evident when jurors were presented with other evidence that favoured guilt.

While jurors were unexpectedly more confident in their guilty verdicts when the eyewitness expressed low confidence rather than high confidence, this effect was qualified with exposure to the unbiased lineup. The counterintuitive main effect suggests that jurors may have

perceived the eyewitness' proclamation of "one-hundred percent" certainty of her identification as unnaturally high. Jurors may have suspected that the eyewitness' confidence level had been artificially inflated by environmental sources, perhaps through witness 'coaching' or rehearsing testimony with a prosecutor. The presence of the unbiased lineup in the qualifying interaction, however, suggests that viewing an unbiased lineup may remove jurors' concern regarding witness coaching. Jurors may have taken note of the perceptual challenge posed to the eyewitness in making a decision from the unbiased lineup, and determined this was a sign of good investigative practice. This appraisal may have dispelled of jurors' concerns of artificial confidence inflation.

Wells et al. (1981) pursued a line of inquiry into witness coaching. In their study, 50% of witnesses were coached by a prosecutor and later testified to jurors. Across all conditions, an average of 34.2% of jurors believed the witness had been coached or rehearsed. However, suspicion that the witness had been coached was unrelated to jurors' decisions to believe the witness or perceived witness confidence. Legal scholars have argued that witness coaching is an unethical practice that "has given trial lawyers a reputation as purveyors of falsehoods" (Luban, 1988 as cited in Gershman, 2002, p. 829). Additional research is needed into how jurors' suspicions of witness coaching are formed.

Expert Testimony and Juror Perceptions

Several studies (as reviewed in Leippe, 1995; Martire & Kemp, 2011) have concluded that expert testimony results in a tendency for jurors to disbelieve eyewitness reports and, consequently, render an increased amount of not guilty verdicts. This pattern did not occur in the present study; expert testimony did not significantly contribute to the predictive ability of the

logistic regression model to predict verdict. Remarkably, jurors rendered verdicts in identical proportions (50.30% not guilty to 49.70% guilty) whether exposed or not exposed to expert testimony This null effect was not due to incomprehension, as expert testimony did not undermine jurors' overall high self-reported ratings of perceived understanding of the trial procedure. In summary, results from this study suggest that jurors understood, but were not directly influenced by, expert testimony in rendering their verdicts.

Similar to findings from Cutler et al. (1989), expert testimony was found to increase the strength of the defence's case. While the researchers designated this as evidence of skepticism, they concluded that that null effects produced by expert testimony in the study on eyewitness credibility, prosecution case strength, and verdict warrant the conclusion that evidence of skepticism was "strikingly small" (p. 222). The present study found an identical pattern of findings. In context, expert testimony should be anticipated to increase the strength of the defence's case from the point of view of the jury. The testimony was presented as defence evidence, and the content of such testimony is designed to address "the potential fallibility of eyewitness identification in the hope that this information will balance the jury's tendency to "overbelieve" eyewitnesses" (Overbeck, 2005, p. 1896). Overall, expert testimony did not result in a tendency for jurors to disbelieve eyewitness testimony.

No evidence was found that expert testimony sensitized jurors to foil bias or eyewitness confidence. While the interpretation of null effects is conventionally unadvisable, this conclusion is supported by findings suggesting that jurors independently and appropriately evaluated the evidence presented in the case in two respects. First, jurors were not sensitized to foil bias because lineup type produced a significant main effect on ratings of lineup fairness, in the

anticipated and appropriate direction. This finding is consistent with previous research and will be discussed. Second, jurors were not sensitized to eyewitness confidence because jurors did not afford undue credibility or probative weight to a highly confident eyewitness. Implications for these findings will be discussed.

Judicial Caution as a Procedural Safeguard

The judicial caution drafted by the Canadian Judicial Council (2012) did not increase the predictive ability of a logistic regression model to predict verdict. The overall trend of split verdicts was evident whether jurors were exposed to the caution (48.4% not guilty to 51.6% guilty) or not exposed to the caution (52% not guilty to 48.0% guilty). Unlike the *Telfaire* instructions explored by Greene (1988), the Canadian model did not induce skepticism in jurors. As hypothesized, the caution did not produce any significant main effects on items in the questionnaire.

The presence of the caution in higher-order interactions indicates that, under specific circumstances, exposure to the judicial caution significantly affected jurors' confidence in their guilty verdicts. Jurors were more confident in their guilty verdicts when not exposed to a lineup, expert testimony, or the caution, compared to jurors who viewed the caution. The caution may have produced a sobering effect on jurors' confidence in the guilt of the defendant, but only compared to a control condition. In the absence of lineup evidence, expert testimony, or the judicial caution, jurors were presented with nothing to challenge their judgments of guilt and rendered high guilty verdict confidence ratings. However, the practical significance of this finding is negligible, as lineups are required to be disclosed in court as evidence under *Stinchcombe*. That 59% of jurors in this condition (n = 39) were willing to convict on the

evidence of a single eyewitness invites research into jurors' criteria of proof of guilt beyond a reasonable doubt in cases resting on eyewitness testimony evidence.

Jurors were more confident in their guilty verdicts when not exposed to expert testimony, exposed to the unbiased lineup, and had viewed the judicial caution compared to jurors who had not viewed the judicial caution. This finding suggests that the caution sensitized jurors to foil bias under specific circumstances, as identifications made from unbiased lineups are more diagnostic of a suspect's guilt (Fitzgerald et al., 2013). A segment of the caution presents a palatable "checklist"-like outline of factors for jurors to consider when evaluating lineup fairness. Upon viewing the caution transcript following exposure to the unbiased lineup, jurors were likely able to determine that the lineup met these guidelines. Affirming the fairness of the lineup may have contributed confirmation bias partial to jurors' beliefs of the suspects' guilt (Nickerson, 1998). Unfortunately, this evidence for sensitization is overshadowed by an identified flaw.

Results from correlational analyses and *t*-tests suggest that exposure to the caution desensitized, or confused, jurors as to the fairness of the eyewitness lineups. Without the judicial caution, jurors rendered lineup fairness ratings consistent with their independent (i.e., absent a procedural safeguard) ability to detect foil bias. When exposed to the judicial caution, fairness ratings from a biased and unbiased lineup no longer differed. The absence of an interaction effect, due to the low statistical power derived for this specific test, does not permit conclusive findings. However, correlational results strongly suggest that jurors' overall ability to detect lineup fairness was hindered with exposure to the judicial caution. Implications for these findings will be addressed.

Disclosing Lineups at Trial

Results from the present study support findings from other studies that have concluded jurors who view lineups can identify the presence of bias (e.g. Devenport et al., 2002; Lindsay, 1994). While Devenport et al. (2002) found that foil bias reduced jurors' ratings of defendant culpability, the present study found that lineup bias has a powerful effect on jurors' verdicts. Lineup type was the only variable to significantly contribute to the ability of a logistic regression model to predict verdict. There was a marked increase in not guilty verdicts from jurors who viewed the biased lineup (61%) compared to jurors who viewed the unbiased lineup (42.1%). It appears jurors decided that the suspect identification from the biased lineup "ought to be treated as weak evidence" (Dupuis & Lindsay, 2007, p. 182).

Results from ANOVAs on items from the Juror Determination Questionnaire offer insight into the increase of not guilty verdicts amongst jurors who viewed the biased lineup. Devenport et al. (2002) hypothesized that foil bias may produce a reverse halo effect on jurors, such that jurors who view a biased lineup may "assume the presence of other biases" in the investigation (p. 1050). The researchers found partial support for this contention through foil bias significantly increasing jurors' ratings of the suggestiveness of the overall identification procedure. Results from the present study support this hypothesis, as exposure to a biased lineup reduced jurors' perceived fairness of the investigation.

The explanation offered above must be interpreted in conjunction with jurors' estimates of the percentage of people who would make an accurate identification in the case. Although jurors who viewed the biased lineup were less likely to convict than jurors who viewed the unbiased lineup, jurors overall did not believe that the biased lineup selection was less likely to

be accurate (except jurors who did not view a safeguard). Introducing the biased lineup as evidence appears to have established reasonable doubt in the minds of jurors due to the lineup being perceived as an unfair investigative practice. Importantly, the increase in not guilty verdicts did not coincide with a decreased estimated percentage of eyewitnesses who would make a correct identification. This explanation must be considered in view of jurors' modest overall estimate of 57.12% of people would make a correct identification.[28]

Limitations

Results of the present study must be viewed in light of important limitations. All SEO designs do not permit a determination of whether procedural safeguards promote jurors' ability to discriminate between accurate and inaccurate eyewitnesses. Studies that employ fictional eyewitnesses, as this one, promote analysis under the framework of juror skepticism, sensitivity, and confusion. As noted by Bornstein (1999), the judiciary is unlikely to consider simulation research in case evaluations. Therefore, SEA studies employing actual eyewitnesses serve as models for researchers looking to objectively compare eyewitness accuracy with jury decision-making, and stand more likely to be referenced in actual cases.

A question of ecological validity is warranted. This study was conducted entirely online, recruited university students, and utilized transcripts as opposed to a live, in-court presentation. While commonly utilized, these study traits have understandably drawn criticism over the generalizability of findings. However, there is evidence that these considerations do not undermine the validity of results. Pezdek, Avila-Mora and Sperry (2009) found no significant

[28] Though the simulated nature of this study disallows postdiction, this estimate approximates actual decision accuracy rates found from meta-analyses of staged-crime studies. Overall, 54.2% of participant eyewitnesses in a meta-analysis of 30 studies by Sporer et al. (1995) made correct identification decisions. Data from participant eyewitnesses included in a study by Steblay, Dysart, Fulero and Lindsay (2003) found that 51% of eyewitnesses made correct identifications.

difference in jurors' perceptions of expert testimony between groups that were exposed to either a transcript or video. Similarly, Bornstein (1999) concluded that different trial presentation mediums, including live, video, audio, written transcript, or written summary, generally produce few differences in juror evaluations. In addition, research suggests that university students render few differences compared to community-sampled jurors (Cutler et al., 1989; Bornstein, 1999).

Logistical constraints placed four further limitations on the ecological validity of this study that are worthy of note. Jurors from PREP pool participated in the minimal trial at their leisure and without supervision. While an effective recruitment method, this approach did not allow the researcher to control the environment or behaviour of participants during completion of the study. This limitation appears inconsequential, however, as no significant differences were found between participant pools. Furthermore, all jurors did not engage in collective deliberations, did not undergo a selection process, and did not observe a 'full' trial. The impact of these limitations is unknown.

Summary

Results suggesting juror confusion from the judicial caution lend important insight to Canadian cases such as *Regina v. Eakin* (1994) that have explicitly acknowledged a caution as sufficient in addressing issues concerning lineup bias, if presented "carefully and thoroughly" (para. 14). The Canadian model was undoubtedly crafted carefully and thoroughly, undergoing major revisions between 2004 and 2012. Evidence in this study supports the conclusion that detecting lineup bias is not outside of the province of the jury. However, their role of triers-of-fact may be better preserved by evaluating eyewitness lineups on their own accord and not exposing them to the caution drafted by the Canadian Judicial Council. Though unintuitive,

results suggest that defence lawyers who are convinced that a lineup used in an investigation was biased towards the suspect should avoid exposing jurors to the Canadian judicial caution. Future studies should continue to investigate jury decision-making across international models of judicial cautions.

Jurors were not influenced by eyewitness confidence in rendering verdicts and did not overbelieve eyewitnesses in general. Based on these findings, it is pertinent to review one purpose of introducing PSs into the courtroom. Where cases involve eyewitness identifications questioned by the defence, PSs educate jurors of the frequently modest CA relationship supported by many studies. Accordingly, jurors are advised to avoid basing determinations of defendant culpability strictly on the confidence of the eyewitness. The findings of the present study put into question the necessity of introducing PSs to meet this specific end. Coupled with the high self-reported understanding of the trial procedure by jurors, a passage from English common law articulates this standpoint: "if on the proven facts a judge or jury can form their own conclusions without help, then the opinion of an expert (or introduction of a caution) is unnecessary" (*Regina v. Turner*, 1975, p. 841 as cited in Martire, 2008).[29]

Jurors rendered a marked increase in not guilty verdicts when exposed to a biased lineup. These results underscore the importance of investigators to evaluate lineup fairness prior to requesting an eyewitness to make an identification decision. As stated by Devenport et al. (2002), "lineup administrators should be made aware of the procedures that make for a good lineup, and these procedures should be implemented nationwide" (p. 1053). In criminal trials launched primarily on eyewitness identification evidence, a biased lineup stands to potentially

[29] Note in parenthesis added.

sabotage conviction efforts of the prosecution – even with a highly confident witness. Combined evidence from Devenport et al. (2002) and the present study provide strong evidence that biased lineups undermine attempts at a conviction and leads jurors to distrust the investigation procedure.

The analysis of the influence of expert testimony lends itself in determining whether eyewitness expert testimony should be admitted into trial under guidelines from *Mohan*. The absence of a direct effect on verdict suggests that jurors upheld their role as a responsible fact-finder and did not accord disproportionate authority to the expert. In addition, expert testimony buttressed the case of the defence but did not undermine the case of the prosecution, reduce estimates of the percentage of people who would make an accurate identification, or decrease jurors' perceived credibility of the eyewitness. These findings indicate that jurors consistently considered the totality of available evidence. However, expert testimony appears unnecessary to assist jurors in assessments of eyewitness confidence and lineup fairness. It is important to note that other studies have found evidence that expert testimony does sensitize jurors to factors not explored in the present study, including weapon focus (Cutler et al, 1989), instruction bias (Devenport et al., 2002), and confidence (Martire & Kemp, 2009; Wells & Wright as cited in Martire & Kemp, 2011). Future studies should employ novel expert testimony presentation methods across unexplored eyewitness identification factors.

Appendix A: Informed Consent Statement (Study 1)

WILFRID LAURIER UNIVERSITY

INFORMED CONSENT STATEMENT

The 'Iterative' Procedure: Lineup Construction and Manipulation in Eyewitness Research

Principal Investigator: Timothy Wykes, MA Candidate

Supervisor: Dr. Jennifer Lavoie, Assistant Professor, Department of Criminology, Laurier Brantford

You are invited to participate in a research study examining evaluations of eyewitness lineups. The purpose of this research is to serve as a preliminary or 'pilot' study to gather summary data from lineup evaluations. Specifically, the lineups that you will view in this stage of the study will be later shown to participant-jurors in a mock-trial setting. You must be 18 years of age or older and a Canadian citizen in order to participate, as these demographics are required in order to participate as a juror in Ontario court trials. You are eligible to participate in this study as a result of your enrollment in an Introductory Psychology course at Wilfrid Laurier University. Accordingly, participation this study is eligible to fulfil partial course credit. Approximately 50 people will participate in this study.

INFORMATION

Should you consent to take part in this research, you will be directed to a webpage that contains instructions on how to proceed with the lineup procedure. Following this, you will receive a general physical description of a 'suspect'. You will then view a photo lineup and be asked to make a selection. This procedure will then repeat for a second lineup. After viewing and selecting a photo from the final lineup, you will then receive a debriefing. It is estimated that your participation will take approximately 10 minutes.

RISKS

Research participants may feel uncomfortable making assessment of individuals in a line-up because they may feel it reveals their own underlying biases. Please remember that you may withdraw from this study at any time. During your participation, you have the right to refuse to answer any question. You may withdraw from the study, or refuse to answer any question, for any reason, without explanation or reason. If you withdraw from the study, your data will be removed from the study and deleted upon request.

Please note that the data will be collected using an online approach, the confidentiality and privacy of data cannot be guaranteed during web transmission.

BENEFITS

Your participation will contribute to an understanding of the selection and evaluation process involved in eyewitness identification. Your responses will also assist in determining the reliability of lineup construction procedures used to create the lineups in this study. Your valued participation will assist the lead researcher in completion of MA Criminology degree requirements at Wilfrid Laurier University.

CONFIDENTIALITY

All information provided for the purposes of this study will be kept strictly confidential. Data accrued in this research study will be summarized and presented in aggregate format. Your valued participation will be used for the purposes of analysis, publication and presentation by the lead researcher. Only the lead researcher and supervisor will have access to the materials accrued in this process, including all responses.

Your responses will be electronically stored on personal and university password-protected computers of the lead researcher and supervisor. In addition, your responses will be held on the hosting website that may be accessed by user name and password combination known only to the lead researcher and supervisor. In the event of the production of physical copies of responses, all materials will be stored in a locked filing unit in the home of the lead researcher.

COMPENSATION

By agreeing to participate, you will fulfil a partial course credit as required under your enrollment in an Introductory Psychology course at Wilfrid Laurier University.

CONTACT

If you have questions at any time about the study or the procedures, or you experience adverse effects as a result of participating in this study you may contact the researcher, Timothy Wykes, at wyke5360@mylaurier.ca. This project has been reviewed and approved by the University Research Ethics Board, REB # 3752. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Dr. Robert Basso, Chair, University Research Ethics Board, Wilfrid Laurier University, (519) 884-1970, extension 4994 or rbasso@wlu.ca.

Please feel free to ask any questions that you may have about this research. If you wish to participate in this study, please indicate as such below.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty. If you

withdraw from the study, your data will be removed and excluded from data analysis upon request. You have the right to omit any question(s)/procedure(s) you choose. The researcher encourages you to cease participation in this study if you see any reason to do so.

FEEDBACK AND PUBLICATION

The results of this research may be disseminated in the form of conference presentations, journal articles, thesis papers and/or book chapters. You may contact Timothy Wykes at any time if you wish to receive information on where to view any disseminations of this research at wyke5360@mylaurier.ca. It is estimated that feedback regarding research results will be available by 30 August, 2014.

Eligibility

○ I am at least 18 years old and a Canadian citizen.

Records

◉ I have read and understand the above information. I have (a) printed a copy of this form for my records and/or (b) understand that I may access this Informed Consent Statement at the current web address on my own time, during data collection, should I wish to review the form for any reason and/or (c) will contact the lead researcher at wyke5360@mylaurier.ca to request a copy. I agree to participate in this study.

Appendix B: Demographic Variables (Study 1)

What is your gender?

- Male
- Female
- Other
- Prefer not to disclose

What is your race? You may select more than one option.

- White/Caucasian
- Black/African-Canadian
- Asian/Asiatic-Canadian
- Indian/Indo-Canadian
- Latino/Spanish
- Aboriginal/First Nations
- Other
- Prefer not to disclose

What is your age in years as of your last birthday?

0 10 20 30 40 50 60 70 80 90 100

Age:

Appendix C: Lineup Procedure Instructions

An eyewitness to a crime has provided a description of the suspect. On the following page, you will read that description. You will then be presented with a set of photographs. You are asked to select the photograph that you believe is the accused. You may select only one photograph. There is no time limit on selection. When you are ready to view the description, proceed to the next page.

Appendix D: Debriefing (Study 1)

The highest known contributing factor to wrongful convictions in the United States is mistaken eyewitness identification. Canada, too, has seen wrongful convictions stemming from mistaken eyewitness identification. Suggestive eyewitness lineups can contribute to mistaken identifications and subsequently, a wrongful conviction.

The first study employing 'mock-witnesses' to evaluate eyewitness lineups dates back to 1973. As you have experienced in this study, this involves receiving a description of a suspect, and being instructed to select from a lineup. If a lineup is fair, no member should stand out. If a lineup is biased, there may be a suspect 'standing out of the crowd'. By collecting these eyewitness selections based on a description, you can evaluate the degree to which a lineup is fair or biased towards a suspect.

The data that you have provided in this study will assist the lead researcher in a follow-up study. Two lineups constructed here will be exposed to jurors in a simulated trial. This second investigation will determine the impact of exposure to eyewitness lineups on juror evaluations. For example: are jurors less likely to render a guilty verdict if an eyewitness to a crime selected the defendant from a biased lineup?

Thank you for participating. Disseminations of this research will be available by 30 August, 2014. You may contact the researcher at wyke5360@mylaurier.ca.

Appendix E: Informed Consent Statement (Study 2)

WILFRID LAURIER UNIVERSITY

INFORMED CONSENT STATEMENT

Juror Perceptions of Evidence in a Criminal Trial

Principal Investigator: Timothy Wykes, MA Candidate
Supervisor: Dr. Jennifer Lavoie, Assistant Professor, Department of Criminology, Laurier Brantford

You are invited to participate in a research study examining how potential jurors perceive and evaluate evidence in a criminal trial. The purpose of this research is to examine if and how case evidence impacts juror perceptions and verdicts. You must be 18 years of age or older, an Ontario citizen, and either (a) enrolled as a student at Laurier Brantford or (b) participating under the Psychology Research Experience Program (PREP) in order to participate. You must be at least 18 and an Ontario citizen as these demographics are required in order to participate as a juror in Ontario court trials. If you are registered in a PREP-eligible psychology course at Wilfrid Laurier University, participation in this study will fulfil partial course credit. It is estimated that approximately 500 people will participate in this study.

INFORMATION

Should you consent to take part in this research, you will be directed to a webpage requesting basic demographic characteristics. You will then view a webpage that contains a summary of information regarding a criminal case. You will be instructed on your role as a juror. You will then view transcripts containing the testimony of numerous people involved in the case, which may include defense and prosecution attorneys, an eyewitness, a judge and/or an expert. You may also view additional evidence submitted in the case. Following the trial, you will be asked to render a verdict (guilty/not guilty) and complete a questionnaire. It is estimated that your participation will take approximately 20-40 minutes.

RISKS

Research participants may feel uncomfortable hearing the testimony of a victim of a robbery. Participants may feel particularly threatened or offended by the witness' recounting of threats made upon her life. If you anticipate that this material will cause any discomfort or offence, the researcher encourages you to decline participation. Please remember that you may withdraw from this study at any time. During your participation, you have the right to refuse to answer any question. You may withdraw from the study, or refuse to answer any question, for any reason, without explanation or reason. If you withdraw from the study, your data will be removed from the study and deleted upon request.

Please note that the data will be collected using an online approach, the confidentiality and privacy of data cannot be guaranteed during web transmission.

BENEFITS

Your participation will contribute to an understanding of the perceptions of jury-eligible citizens of Ontario in cases that rely primarily on identification evidence. Students enrolled in eligible psychology courses at Wilfrid Laurier University will fulfill partial course credit through participation. Your valued participation will assist the lead researcher in completion of MA Criminology degree requirements at Wilfrid Laurier University.

CONFIDENTIALITY

All information provided for the purposes of this study will be kept strictly confidential. Your valued participation will be used for the purposes of analysis, publication and presentation by the lead researcher. Only the lead researcher and supervisor will have access to the materials accrued in this process, including all responses. If you are enrolled in an eligible psychology course and choose to participate for partial course credit, your name will be collected in order for the researcher to grant credit. None of your personal details or responses will be shared with third parties.

Your responses will be electronically stored on personal and university password-protected computers of the lead researcher and supervisor. In addition, your responses will be held on the hosting website that may be accessed by user name and password combination known only to the lead researcher and supervisor. In the event of the production of physical copies of responses, all materials will be stored in a locked filing unit in the home of the lead researcher.

The researcher may wish to use direct quotes from your responses to open-ended questions. These quotes may be used in any analysis, publication or presentation by the lead researcher. Quotations will not be attributed to you. If you do not wish to have your exact words quoted (with the safeguard of no name attribution) you may decline below and still participate. In the event that quotes contain potentially identifying information, the researcher will alter and/or remove that information.

COMPENSATION

Eligible PREP students will fulfil partial course credit as required under enrollment at Wilfrid Laurier University.

Non-PREP student participants will participate strictly voluntarily and without compensation.

CONTACT

If you have questions at any time about the study or the procedures, or you experience adverse effects as a result of participating in this study you may contact the researcher, Timothy Wykes, at wyke5360@mylaurier.ca. This project has been reviewed and approved by the University Research Ethics Board, REB # 3798. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Dr. Robert Basso, Chair, University Research Ethics Board, Wilfrid Laurier University, (519) 884-1970, extension 4994 or rbasso@wlu.ca.

Please feel free to ask any questions that you may have about this research. If you wish to participate in this study, please indicate as such below.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty. If you withdraw from the study, your data will be removed and excluded from data analysis upon request. You have the right to omit any question(s)/procedure(s) you choose. The researcher encourages you to cease participation in this study if you see any reason to do so.

FEEDBACK AND PUBLICATION

The results of this research may be disseminated in the form of conference presentations, journal articles, thesis papers and/or book chapters. You may contact Timothy Wykes at any time if you wish to receive information on where to view any disseminations of this research at wyke5360@mylaurier.ca. It is estimated that feedback regarding research results will be available by 30 August, 2014.

I am at least 18 years of age and a citizen of Ontario, Canada.

○ Yes

○ No

I have read and understand the above information. I have (a) printed a copy of this form for my records *and/or* (b) understand that I may access this Informed Consent Statement at the current web address on my own time, during data collection, should I wish to review the form for any reason *and/or* (c) will contact the lead researcher at wyke5360@mylaurier.ca to request a copy. I agree to participate in this study.

○ Yes

○ No

(Optional) I agree to allow direct (anonymous) quotations in publications.

- Yes
- No

(Optional) The researcher may contact me in the future, if necessary, to ask for clarification, or elaboration, with respect to information that I provide.

- Yes
- No

Appendix F: Juror Demographics (Study 2)

What is your gender?

- Male
- Female
- Other
- Prefer not to disclose

What is your race?

- White/Caucasian
- Black
- Asian
- Indian
- Latino/Spanish
- Aboriginal/First Nations
- Biracial/Multiracial
- Other
- Prefer not to disclose

What is your age in years as of your last birthday? (Numerical format; write 'NA' if you prefer not to disclose).

Please indicate your eligibility for this study. If you are participating in this study for partial course credit under PREP, please select option A, regardless of your campus enrollment.

- Psychology Research Experience Program (PREP)
- Laurier Brantford student

Appendix G: Case Facts

You are a juror in a criminal trial. The accused, Ronald Carns, has pleaded not guilty to the charge of robbery. Under s. 343(a) of the Canadian Criminal Code, every one commits robbery who: (a) *steals, and for the purpose of extorting whatever is stolen or to prevent or overcome resistance to the stealing, uses violence or threats of violence to a person or property*. Under s. 344, every person who commits robbery is guilty of an indictable offence and liable to imprisonment for life.

The case against Mr. Carns rests primarily on a sole eyewitness. The eyewitness will discuss general circumstances surrounding the robbery. You will read that Ms. Koswick selected the defendant out of a photographic lineup. This identification was the primary reason charges were laid against the defendant.

You will read testimony from people involved in the case. Pay *careful attention* to the testimony. As a juror, you will be asked to render a verdict of guilty or not guilty following the trial. These proceedings are based on actual criminal events. It is vital that you render your verdict and answer all questions as if you were sitting in an *authentic court trial*. If for any reason you are unable to commit to this role, please remove yourself from this study now[30].

You are encouraged to take notes.

[30] Underlined portion provided an exit link.

Appendix H: Eyewitness Direct Examination

DEFENSE: Good afternoon.
EYEWITNESS: Hi.
DEFENSE: Can you state your name for the record?
EYEWITNESS: Jamie, J-A-M-I-E, Koswick, K-O-S-W-I-C-K.
DEFENSE: What is your date of birth, Ms. Koswick?
EYEWITNESS: September 20th, 1984.
DEFENSE: Are you employed?
EYEWITNESS: Yes, at Pioneer.
DEFENSE: The convenience store?
EYEWITNESS: Yes.
DEFENSE: Where is that located?
EYEWITNESS: In Toronto. The Beaches area.
DEFENSE: How long have you been there?
EYEWITNESS: Almost 2 years.
DEFENSE: Okay. What do you do there?
EYEWITNESS: I'm the Assistant Manager, so I take care of orders, stocking, schedules, clerical duties…
DEFENSE: Okay. I just want to ask you a few questions about your day at work. Before the event.
EYEWITNESS: Okay.
DEFENSE: First, can you tell the ladies and gentlemen of the jury what day the event in question happened?
EYEWITNESS: It was August 13th, 2012.
DEFENSE: Was that a Monday?
EYEWITNESS: Yes.
DEFENSE: Okay. And what time did you get in?
EYEWITNESS: Just before 6:00 AM. I was opening.
DEFENSE: And what time were you scheduled until that day?
EYEWITNESS: 5:00 PM. I had someone coming in to take over at 5:00 PM.
DEFENSE: Okay. What happened during that day? Before the event?
EYEWITNESS: I mostly cleaned. Re-stocked. It was uneventful. I don't think we had a lot of customers that day, because gas was up, and it was busy the day before.
DEFENSE: Gas was up?
EYEWITNESS: In price. It went up overnight.
DEFENSE: Okay. Anything else you remember before the event during the day of August 13th, 2012? Anything out of the ordinary happen?
EYEWITNESS: No, not really. It was warm in the store. Nothing else important that I remember.
DEFENSE: Okay. I'm going to turn to the event in question now. When did you notice the person who robbed you walked in the door?
EYEWITNESS: I heard the 'dings' over at the door at about a quarter after 2. I looked up when I heard the noise. So I first saw him as soon as he walked in.

DEFENSE: Can you just clarify what the 'ding' is?
EYEWITNESS: Oh. Sure, okay. It's just the sound the door makes –
DEFENSE: When someone walks in?
EYEWITNESS: Yeah. Yes.
DEFENSE: Okay. And what did you see?
EYEWITNESS: I saw a man…he was just walking in and I could see him from the counter.
DEFENSE: Did you recognize him?
EYEWITNESS: I didn't ever recall seeing him before. So, no. As far as I could tell he was a stranger.
DEFENSE: Okay. And what happened next?
EYEWITNESS: I looked at him walking in. He was walking really quickly and he went right towards the counter.
DEFENSE: Walking quickly? How?
EYEWITNESS: He went right from the door into right – right ahead of the counter. Straight ahead, and walked right up to me.
DEFENSE: And he was walking quickly?
EYEWITNESS: Yes. He was rushing.
DEFENSE: What else?
EYEWITNESS: He was looking around. There was no one else in the store or at the pumps. He looked around before he came to the counter and said, "Open it. Give me all the money now."
DEFENSE: "Give me all the money now" – did he say anything else?
EYEWITNESS: No. It was just quiet at first…"give me the money…" and he got louder: "give me the money now!"
DEFENSE: How were you feeling at this point?
EYEWITNESS: I stayed as calm as I could.
DEFENSE: How did you do that?
EYEWITNESS: I had been trained about what to do when there's a robbery. Gas stations get robbed a lot so all managers and assistant managers have been trained on how to deal with it. We're told to just stay calm. So I just opened the register.
DEFENSE: Okay. So did you do what you were trained to do?
EYEWITNESS: Actually, yeah. I opened the register.
DEFENSE: Okay. What happened next?
EYEWITNESS: I was going to hand over bills but he screamed "back up, back up…I'll kill you" so I went against the back of the counter. Put my hands up to show him I wasn't pushing any alarms or anything. He said it twice. Back-to-back: "I'll kill you…I will kill you." Emphasizing it a lot.
DEFENSE: What did he do after you moved back?
EYEWITNESS: He reached over the counter and pulled out the bills, out of the register. He got almost all of them. A few dropped. I just stood there and watched him.
DEFENSE: Okay, can you describe, for the jury, how he reached over?
EYEWITNESS: He just used his right hand and just stuffed the money into his pockets.
DEFENSE: No change?
EYEWITNESS: No, just bills.
DEFENSE: Let me back up a bit - how did you open the register?

EYEWITNESS: The key was already in it, I just turned it to open it.
DEFENSE: It's not one of those registers that are really hard to open without making a sale?
EYEWITNESS: No, it's very easy to open if you just leave the key in it.
DEFENSE: Okay, but he took the bills out?
EYEWITNESS: Yes.
DEFENSE: And did you see his face?
EYEWITNESS: Yes, when I was not opening the register, I looked at his face.
DEFENSE: Okay. Then what?
EYEWITNESS: He just ran out.
DEFENSE: How did he run out?
EYEWITNESS: He turned around, and just ran. Yeah, and went out the door and made a left.
DEFENSE: But you saw his face before this…
EYEWITNESS: Yeah – yes. He was facing me during.
DEFENSE: How good of look did you get at his face?
EYEWITNESS: It was fine…bright. Lights were all working and everything, and it was the day.
DEFENSE: So it was a good look.
EYEWITNESS: Yeah, it was clear.
DEFENSE: Okay. What did you do after he left?
EYEWITNESS: I pushed the panic alarm. It's underneath the counter so I couldn't get to it when I was standing back. The police came very quickly after 911 called me and asked if I had been robbed.
DEFENSE: Did the police attend?
EYEWITNESS: Yes, pretty fast.
DEFENSE: What did you tell 911?
EYEWITNESS: Just what I said today, basically. I described him. I said which way he went.
DEFENSE: Does your store have security cameras?
EYEWITNESS: Our security cameras weren't working. We had some sort of electric issue….we think one of our delivery guys accidentally severed a wire when we got our delivery on the Thursday of the week before. But we hadn't checked the cameras until the police asked for them…so we didn't know. Nothing was recorded, everything was off.
DEFENSE: So there is no footage of the robbery?
EYEWITNESS: Not that I know of. No.
DEFENSE: And there's no one else that can back-up your testimony?
EYEWITNESS: I was alone in the store, there was no one at the pumps…I don't know of anyone that could have seen anything else. I think if there was, they'd be here.
DEFENSE: Okay. Do you see the person that robbed you today?
EYEWITNESS: Yes, I do.
DEFENSE: Can you point him out? Let the record show that the witness pointed at the defendant, Ronald Carns. Okay. Now, I need to fast forward a bit – to talk about your time at the police station.
EYEWITNESS: The detective in charge of the case gave me a call and I came in.
DEFENSE: And what happened when you attended the police station?
EYEWITNESS: I drove to the station and I met the detective. He asked me to come into a room and view a bunch of photos.

DEFENSE: Okay. So what happened? Did you recognize anyone?
EYEWITNESS: Well, I came to one, and I looked at the photo and said 'him', and pointed to the photo…
DEFENSE: I'd like to turn the jury's attention to the copy of the photo lineup that was used. Ladies and gentlemen of the jury, if you would, please have a look at the photo lineup that Ms. Koswick viewed at the station. Ms. Koswick, did you pick out the person that is circled?
EYEWITNESS: Yes.
DEFENSE: Let the record show that she made a positive identification of the defendant, as indicated. Ms. Koswick, I'm almost done. I just want to ask, how sure are you that it was Ronald Carns that robbed you that day?
EYEWITNESS: (*High Confidence*): I'm one-hundred percent sure. Positive. (*Low Confidence*): I'm pretty sure it was him.
DEFENSE: No further questions, your Honour.

Proceed

Appendix I: Eyewitness Cross-Examination

DEFENSE: Hi, Ms. Koswick. I'm not going to take up too much of your time, okay? So, you are confident that the person you identified is the same person you saw at the robbery? You want the jury to believe that?
EYEWITNESS: Yes.
DEFENSE: You picked the defendant out of that lineup - the one admitted as evidence? And you're positive?
EYEWITNESS: Yes.
DEFENSE: And today how do you feel?
EYEWITNESS: (*High Confidence)*: Like I said, I'm one-hundred percent sure it was him. (*Low Confidence)*: Well, like I said, I'm pretty sure it was him.
DEFENSE: Did you ever think you may have chosen the wrong person? That Mr. Carns is an innocent man? That he is not the person that robbed you?
EYEWITNESS: (*High Confidence)*: Not at all. (*Low Confidence)*: No…
DEFENSE: Ladies and gentlemen, I encourage you to think hard about this identification. Thank you, your Honour. No further questions.

Proceed

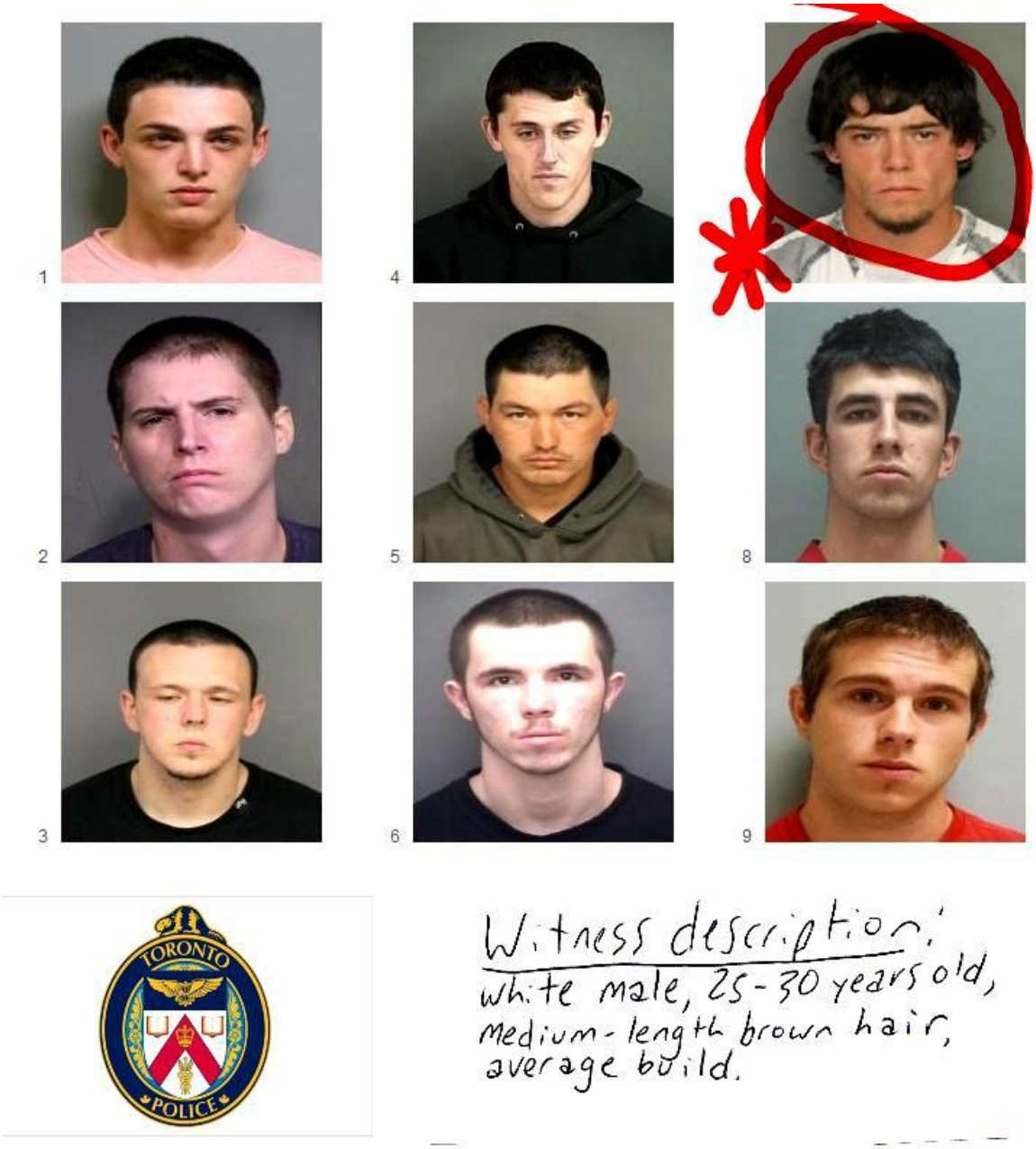
Appendix J: Unbiased Lineup

Study the evidence and select 'Proceed' to continue[31].

Proceed

[31] This lineup is the same as was evaluated in Study 1. For Study 2, the red indications of the suspect, police service logo, and handwritten witness description were added.

Appendix K: Biased Lineup



Study the evidence and select 'Proceed' to continue[32].

Proceed

[32] See above, note 28.

Appendix L: Expert Witness Direct Examination

DEFENSE: Good afternoon.
EXPERT: Good afternoon, counsel.
DEFENSE: Would you tell the ladies and gentlemen of the jury your name, please?
EXPERT: My name is Dr. Sam Barnes.
DEFENSE: Dr. Barnes, your doctorate, Ph.D., is in psychology – is that correct?
EXPERT: Correct.
DEFENSE: And you're aware that we, the defense, have chosen not to admit any other evidence today besides your testimony?
EXPERT: Yes.
DEFENSE: Okay. Will you tell the ladies and gentlemen of the jury a little bit about your qualifications? First, your education.
EXPERT: I was an undergraduate at University of Toronto. I received a bachelor's degree in psychology with a minor in law. It was awarded in 1972. I moved to British Columbia to attend Simon Fraser University in 1973. I received a master's degree in psychology in 1974, followed three years later by a Ph.D. in psychology in 1977.
DEFENSE: What field of research are you generally involved with?
EXPERT: Within the field of psychology, my major specialty is human memory, eyewitness testimony, and other topics related to the subject of memory.
DEFENSE: What teaching experience do you have?
EXPERT: I taught various psychology courses at Simon Fraser towards the completion of my Ph.D. I was signed first as a contract faculty member at Queen's University in Kingston and then accepted a full-time tenure track position in the department of psychology. I currently teach graduate level courses in memory. I am also a visiting lecturer at Stanford University in the US.
DEFENSE: Can you tell the jury about the studies that researchers such as yourself do?
EXPERT: Eyewitness researchers collect scientific information using experimental studies. We conduct experiments to understand what factors can influence identification accuracy. We look for overall trends in research to determine with a degree of certainty what factors consistently can help an eyewitness be accurate, and what can cause them to make mistakes. We also analyze known cases of wrongful convictions. After a convicted person has been exonerated by DNA evidence, we can revisit why people were wrongfully convicted. The majority of these cases you see some problem with eyewitness testimony as a major cause of wrongful conviction. These have been very helpful.
DEFENSE: Provincial governments have gotten involved in publishing on miscarriages of justice like those, is that right?
EXPERT: That's correct. The best known is probably the Inquiry into the case of Thomas Sophonow, a man who was wrongfully convicted primarily through eyewitness misidentification. It included a lot of suggestions to change the process of collecting eyewitness evidence and conducting lineups.
DEFENSE: Are reports like these actually saying: We found that these are bad prosecutions because of these reasons? How many are there?
EXPERT: They were certainly mistaken prosecutions or mistaken outcomes. There have been over 300 confirmed wrongful convictions in the United States, and several in Canada.

DEFENSE: Okay. Now, you know that main evidence in this case is the testimony of a single eyewitness?
EXPERT: I do, yes.
DEFENSE: And you cannot legally comment on the actual lineup in this case.
EXPERT: That's correct.
DEFENSE: Okay. You've testified frequently, haven't you?
EXPERT: I have over the last -- since 1990, maybe 80 times.
DEFENSE: Is there a reason for that?
EXPERT: Eyewitness evidence can be very persuasive testimony, even when it's wrong or even when it's got some serious problems. Usually my research focuses on what some of those problems with eyewitness testimony might be under certain conditions.
DEFENSE: Now, you're not going to testify in this case and you don't testify in cases that somebody didn't see what they said they saw.
EXPERT: No. For the most part, no, you can't say whether the witness was right or wrong, only that certain factors may be present that lead an average person to make a mistake.
DEFENSE: And you're not here to testify that someone's right or wrong, but only to offer your expertise on what might affect the accuracy of eyewitness testimony?
EXPERT: Absolutely. Yes.
DEFENSE: Tell us generally, then, what are the things that we should look for. What are the things that affect the reliability of someone who says, "I saw something." What are the things that affect how the jury should judge that?
EXPERT: Well, the first thing you need to realize is that memory does not work like a videotape recorder. We do not remember things perfectly; we don't record an event and play it back later. The process is much more complex. We divide the process into three major stages. So, first there is the *acquisition stage*. Here's a period of time where some event occurs and some information in laid down in a witness' memory system. Then the event is over, time is passing, and we refer to this second stage as the *retention stage*. And finally, you might approach a witness and ask certain questions or ask the witness to make an identification, ask the witness to retrieve information from memory. We call this the *retrieval stage*. Our job as researchers in this area then is to identify the factors that come into play in each of these three stages that affect the accuracy of someone's recollection.
DEFENSE: If a person is shown, for instance, a lineup that is not put together very well so that the suspected person resembles the description of the culprit but none of the others in the lineup resemble the culprit, can that alter the memory?
EXPERT: That would be a biased test. So, if you show an eyewitness some photos, it might be in a booklet form. The witness is asked to pick out the person that they think committed the crime. If there is one person in that photo set who matches a prior description given by an eyewitness, then many witnesses will often fasten on that one individual whether that person is the right person or the wrong person. And, of course, once that selection is made, it then influences everything that happens after that.
DEFENSE: How can this be avoided? How should photo lineups look?
EXPERT: The best answer I can give is that it is of the utmost importance for every person to have the same general appearance. That is, everyone in the lineup should have the same basic characteristics that were described by the eyewitness making the identification. Ensuring that the

lineup is ‘fair’ assists in ensuring that an identification is being made by a witness with a clear memory. Otherwise, the witness may be relying on what is known as relative judgments – simply selecting a person based on the fact that they match the description. So, your goal is to try to create a test, whether it's a photographic test or whether it's a live lineup in which all people fit the description of the perpetrator.
DEFENSE: What about when someone says they are positive that they are right? That they selected the right person?
EXPERT: Well, when you ask that question, you're asking about a very important factor that many experts have studied and testified about, and that’s confidence. When somebody, say, makes an identification or makes a retrieval, tries to answer a question based on their memory, they often will say how confident they are when they do that. And one of the things that the studies have shown is the relationship between how confident somebody is about their identification and how accurate they are is actually weak relationship overall. A person on the street thinks it's a strong relationship, but the scientific work in the area shows it can be a weak relationship. It's common to have people who are very confident and wrong. But, of course, other times people who are very confident are accurate.
DEFENSE: The incorrect person wouldn't be lying, but actually believes what she or he is saying?
EXPERT: Oh, absolutely. People do come to believe that inaccurate identifications are from their own recollection, and they are not intentionally lying.
DEFENSE: And there are others, many others that agree with your work? Professionals, I mean.
EXPERT: Absolutely, yes.
DEFENSE COUNSEL: No further questions at this time. We pass the witness to the prosecution for cross-examination.

Appendix M: Expert Witness Cross-Examination

CROWN: Good afternoon, Dr. Barnes.
EXPERT: Good afternoon.
CROWN: My name is Seth Bowlen, I am the prosecuting attorney in this case. Basically what you testify about is, it's not really hard science, it's more soft science?
EXPERT: We – social scientists – testify around general social frameworks. These inferences are drawn from memory and eyewitness research. There are established trends and developments, from which general conclusions about the accuracy of memory under certain conditions can be made.
CROWN: But these trends are not laws, now are they?
EXPERT: Correct.
CROWN: Okay. Are some of your fellow psychologists in disagreement about the appropriateness of expert testimony?
EXPERT: Yes.
CROWN: Do other psychologists and psychiatrists argue that your research is unproven in real life situations?
EXPERT: Some do, yes.
CROWN: You can never exactly recreate a real-life traumatic event like a violent robbery, like this case, can you?
EXPERT: Well, not ethically, no. That's why we have to look at actual cases and also do the experimental work and look at the collection of data together.
CROWN: And the experimental work basically boils down to you having some of your university students come in and watch a film and then, after the film, you ask them a bunch of questions.
EXPERT: That happens in some of the studies. We do stage robberies in films and have used violent acts in the films.
CROWN: But only in films?
EXPERT: Violent acts like that, yes.
CROWN: So, there's no way you can know for sure what witnesses would do in a real-life situation because you can't scientifically recreate that type of event?
EXPERT: Well, we certainly can't subject witnesses in the experiments to, you know, a whole lot of trauma that might harm them. That's why we have to do what we can in the experimental studies and complement the analysis with actual analyses, with real-life cases.
CROWN: And the people who criticize your experiments are other psychologists?
EXPERT: Sometimes, yes.
CROWN: Okay. I know that you cannot comment on the photos in the case, but, talking about lineups for a minute. Defense counsel asked you some questions about the lineups in this case. When you think of suggestive or biased lineups, it may be where they had a black suspect in the lineup with five white men. That's a suggestive lineup?
EXPERT: Yes.
CROWN: Or where the offender was in his teens and he was in the lineup with all people over the age of 40. That's a suggestive lineup?
EXPERT: Yes. Those are extreme examples.

CROWN: Now, you're mainly here because of the photo lineup. And you spoke about all of these different ways to make a lineup of suspects. But if someone points out a suspect from even an unfair lineup, they could still be correct, right?
EXPERT: Correct.
CROWN: Okay. One last thing. You mentioned that people that are more confident when they pick someone are not necessarily more accurate.
EXPERT: What has been found is that the overall relationship between confidence and accuracy is weak.
CROWN: But isn't it true that people who pick a person from a lineup – choosers – have a better relationship with accuracy? Isn't the fact that someone chose a person a more meaningful predictor of accuracy than someone who says, "I don't see the person"?
EXPERT: Yes, that is consistent with the research. Accuracy is generally positively related to confidence when an eyewitness makes a choice.
CROWN: Thank you, Dr. Barnes. No further questions at this time.

Proceed

Appendix N: Final Judicial Instructions

Identification is an important issue in this case. The case against Mr. Carns depends entirely (or to a large extent) on eyewitness testimony.

You must be very careful about relying on eyewitness testimony to find Mr. Carns guilty of any criminal offence. Innocent people have been wrongly convicted because reliance was placed on mistaken eyewitness identification. Even a number of witnesses can be honestly mistaken about identification. Eyewitness identification may seem more reliable than it actually is because it comes from a credible and convincing witness who honestly but mistakenly believes that the accused person is the one he or she saw committing the offence.

There is little connection between great confidence of the witness and the accuracy of the identification. Even a very confident witness may be honestly mistaken. A very confident witness may be entirely wrong with respect to his or her identification evidence.

Eyewitness identification is a conclusion based on the witness's observations. The reliability of the identification depends on the basis for the witness's conclusion.

Consider the various factors that relate specifically to the eyewitness and her identification of Mr. Carns as the person who committed the offence charged:

The Court asks you to consider the circumstances in which the witness made her observations. Does the witness have a reliable memory? The witness provided the description immediately after the robbery on 911. You have heard that the witness did not recognize the suspect. The witness has testified that visibility was clear and the environment was well-lit. The witness has testified that suspect and the witness had some time of relatively close proximity. Please consider the description provided by the witness after she made her observations. Consider how specific the description was and how close the description is to the way Mr. Carns looked.

As for the lineup, photographs of others were also shown. Consider what the witness said when she identified Mr. Carns from the lineup, with the detective present. The witness has not failed to identify Mr. Carns. The witness has not changed her mind about the identification. Consider: Was the line-up procedure fair? Did the other participants in the line-up share the physical characteristics of the accused? Were the photos similar - for example, in size and colour?

Remember, the Crown must prove beyond a reasonable doubt that it was Mr. Ronald Carns who committed the offense charged. Consider the evidence of the identification witness along with the other evidence you have seen and heard in deciding that question.

Appendix O: Reasonable Doubt

The principle of "proof beyond a reasonable doubt" is an essential part of the presumption of innocence. A reasonable doubt is not a far-fetched or frivolous doubt. It is not a doubt based on sympathy or prejudice. It is a doubt based on reason and common sense. It is a doubt that arises at the end of the case based not only on what the evidence tells you but also on what that evidence does not tell you.

It is not enough for you to believe that Mr. Carns is probably or likely guilty. In those circumstances, you must find him/her not guilty, because the Crown would have failed to prove his/her guilt beyond a reasonable doubt. Proof of probable or likely guilt is not proof of guilt beyond a reasonable doubt.

You should also remember, however, that it is nearly impossible to prove anything with absolute certainty. The Crown is not required to do so. Absolute certainty is a standard of proof that does not exist in law.

If, at the end of the case, and after an assessment of all of the evidence, you are not sure that Mr. Carns committed the offence of robbery, you must find him/her not guilty.

If, at the end of the case, based on all of the evidence, you are sure that Mr. Carns committed the offence of robbery, you should find Mr. Carns guilty.

 Proceed

Appendix P: Juror Determination Questionnaire

What is your verdict in this case?

- ○ Guilty
- ○ Not Guilty

How confident are you in your verdict? (1 = Not at all confident, 9 = Completely confident)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Rating:	○	○	○	○	○	○	○	○	○

How credible did you find the eyewitness in the case? (1 = Not at all credible, 9 = Completely credible)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Rating:	◉	○	○	○	○	○	○	○	○

How confident was the eyewitness in this case? (1 = Not at all confident, 9 = Completely confident)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Rating:	○	○	○	○	○	○	○	○	○

How fair was the lineup in this case? (1 = Not at all fair, 9 = Completely fair)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Rating:	○	○	○	○	○	○	○	○	○

How well did you understand the trial procedure? (1 = Did not at all understand, 9 = Completely understood)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Rating:	○	○	○	○	○	○	○	○	○

Overall, did the investigating officers in this case conduct a fair investigation? (1 = Not at all fair, 9 = Completely fair)

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

	1	2	3	4	5	6	7	8	9
Rating:	○	○	○	○	○	○	○	○	○

How strong is the prosecution's case? (1 = Not at all strong, 9 = Completely strong)

	1	2	3	4	5	6	7	8	9
Rating:	○	○	○	○	○	○	○	○	○

How strong is the defense's case? (1 = Not at all strong, 9 = Completely strong)

	1	2	3	4	5	6	7	8	9
Rating:	○	○	○	○	○	○	○	○	○

What percentage of people do you think would make an accurate identification in this case? (0 - 100%)

0 10 20 30 40 50 60 70 80 90 100

Rating (sliding scale 0 – 100): _____

(Optional) Copy and paste the notes you took during the trial here:

Appendix Q: Debriefing (Study 2)

Canada and the United States have seen wrongful convictions stemming from mistaken eyewitness identification. Researchers have found that suggestive lineups can contribute to mistaken identifications and subsequently, a wrongful conviction. Canada has enacted safeguards to assist in how jurors evaluate eyewitness evidence. Accordingly, researchers have evaluated these safeguards by running experimental trials.

As you have experienced in this study, many factors must be considered by jurors weighing eyewitness testimony. Your participation has lent insight into the efficacy of such safeguards and into how evidence impacts verdicts.

Thank you for participating. Disseminations of this research will be available by 30 August, 2014. You may contact the researcher at wyke5360@mylaurier.ca.

References

Association in Defense of the Wrongfully Convicted. (2014). *Historical cases*. Retrieved from http://www.aidwyc.org/cases/historical/

Agresti, A., & Agresti, B. F. (1978). Statistical analysis of qualitative variation. *Sociological Methodology*, *9*, 204-237.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Beaudry, J. L., Lindsay, R. C., Leach, A. M., Mansour, J. K., Bertrand, M. I., & Kalmet, N. (2013). The effect of evidence type, identification accuracy, line-up presentation, and line-up administration on observers' perceptions of eyewitnesses. *Legal and Criminological Psychology*, DOI: 10.1111/lcrp.12030.

Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychology, 20*(1), 115-129.

Boneau, C. A. (1960). The effects of violations of assumptions underlying the *t* test. *Psychological Bulletin, 57*(1), 49-64.

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, *23*(1), 75-92.

Bornstein, B. & Hamm, J. A. (2012). Jury instructions on witness identification. *Court Review: The Journal of the American Judges Association*, Paper 377, retrieved from http://www.digitalcommuns.unl.edu/adjcourtreview/377.

Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence-accuracy relationship in eyewitness identification: the effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied, 8*(1), 44-56.

Brewer, N. & Wells, G.L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*, 11–30.

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

Brigham, J. C., Meissner, C. A., & Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology, 13*, S73-S92.

Bromby, M., MacMillan, M., & McKellar, P. (2007). An examination of criminal jury directions in relation to eyewitness identification in commonwealth jurisdictions. *Common Law World Review, 36*(4), 303-336.

Buckhout, R. (1976). Eyewitness testimony. In R. Held & W. Richards (Eds.), *Recent progress in perception*. San Francisco: Freeman.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review, 7*(1), 26-48.

Canadian Judicial Council (2012). *Judicial instructions – The National Judicial Institute*. Retrieved from https://www.nji-inm.ca/nji/inm/progs-res/jury-instructions-les-directives-au-jury.cfm?lang=en.

Charman, S. D., & Cahill, B. S. (2012). Witnesses' memories for lineup fillers postdicts their identification accuracy. *Journal of Applied Research in Memory and Cognition, 1*(1), 11-17.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement, 33*(1), 107-112.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.) Hillside, NJ: Erlbaum.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics, 19*(1), 15-19.

Cook, R. D. & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman & Hall.

Cory, P. D. C. (2011). *The inquiry regarding Thomas Sophonow: The investigation, prosecution and consideration of entitlement to compensation*. Province of Manitoba: Ministry of the Attorney General.

Corey, D., Malpass, R. S., & McQuiston, D. E. (1999). Parallelism in eyewitness and mock witness identifications. *Applied Cognitive Psychology, 13*, S41-S58.

Cox, D. R. & Snell, D. J. (1989). *The analysis of binary data* (2nd ed.). London: Chapman & Hall.

Crichton, N. (2001). Wald test. *Journal of Clinical Nursing, 10*, 774.

Cutler, B. L., Dexter, H. R., & Penrod, S. D. (1989). Expert testimony and jury decision making: An empirical analysis. *Behavioral Sciences & the Law, 7*(2), 215-225.

Cutler, B. L., Dexter, H. R., & Penrod, S. D. (1990). Nonadversarial methods for sensitizing jurors to eyewitness evidence. *Journal of Applied Social Psychology, 20*(14), 1197-1207.

Cutler, B. L. and Kovera, M. B. (2010). *Evaluating eyewitness identification*. New York, NY: Oxford.

Cutler, B. L. & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology, and the law.* New York, NY: Cambridge.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1989). The eyewitness, the expert psychologist, and the jury. *Law and Human Behavior, 13*(3), 311-332.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11*(3), 233-258.

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

Daubert v. Merrell Dow Pharmaceuticals, Inc, 113 S. Ct. 2786 (1993).

Devenport, J. L., & Cutler, B. L. (2004). Impact of defense-only and opposing eyewitness experts on juror judgments. *Law and Human Behavior, 28*(5), 569-576.

Devenport, J. L., Penrod, S. D., & Cutler, B. L. (1997). Eyewitness identification evidence: Evaluating commonsense evaluations. *Psychology, Public Policy, and Law, 3*(2-3), 338-361.

Devenport, J. L., Stinson, V., Cutler, B. L. and Kravitz, D. A. (2002). How effective are the expert testimony and cross-examination safeguards? Jurors' perceptions of the suggestiveness and fairness of biased lineup procedures. *Journal of Applied Psychology, 87*, 1042-1054.

Doob, A. N. and Kirshenbaum, H. M. (1973). Bias in police lineups ± partial remembering. *Journal of Police Science and Administration, 1*, 287-293.

Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics, 19*(1), 61-74.

Dupuis, P.R. & Lindsay, R.C.L. (2007). Radical alternatives to traditional lineups. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia, *The handbook of eyewitness psychology (Vol. II): Memory for people* (pp. 179 – 200). Mahwah , NJ: Lawrence Erlbaum & Associates.

Fanselow, M. (1975). How to bias an eyewitness. *Social Action and the Law, 2*(3), 3-4.

Federal Evidence Review (2014). *Federal Rules of Evidence 2014*. Retrieved from http://federalevidence.com/downloads/rules.of.evidence.pdf

Federal Youth Corrections Act, ch. 1115, 64 Stat. 1085 (1950).

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.) Thousand Oaks, CA: Sage.

Fitzgerald, R. J., Price, H. L., Oriet, C., & Charman, S. D. (2013). The effect of suspect-filler similarity on eyewitness identification decisions: A meta-analysis. *Psychology, Public Policy, and Law, 19*(2), 151-164.

Fox, S. G. and Walters, H. A. (1986). The impact of general versus specific expert testimony and eyewitness confidence upon mock juror judgment. *Law and Human Behavior, 10*, 215-228.

Fradella, H. F. (2006). Why judges should admit expert testimony on the unreliability of eyewitness testimony. *Federal Courts Law Review, 3*, 1-29.

Frye v. United States, 293 F.1013 (D.C. Cir. 1923)

Gershman, B. L. (2001). Witness coaching by prosecutors. *Cardozo Law Review, 23*, 829-862.

Garrioch, L., & Brimacombe, C. E. (2001). Lineup administrators' expectations: Their impact on eyewitness confidence. *Law and Human Behavior, 25*(3), 299-315.

Gonzalez, R., Ellsworth, P. C., & Pembroke, M. (1993). Response biases in lineups and showups. *Journal of Personality and Social Psychology*, *64*(4), 525-537.

Greene, E. (1988). Judge's instruction on eyewitness testimony: Evaluation and revision. *Journal of Applied Social Psychology, 18*(3), 252-276.

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using receiver operating characteristic analysis. *Current Directions in Psychological Science*, *23*(1), 3-10.

Haw, R. M., & Fisher, R. P. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology*, *89*(6), 1106-1112.

Hosch, H. M., Beck, E. L., & McIntyre, P. (1980). Influence of expert testimony regarding eyewitness accuracy on jury decisions. *Law and Human Behavior*, *4*(4), 287-296.

Hosmer Jr, D. W., & Lemeshow, S. (2004). *Applied logistic regression*. New York: John Wiley & Sons.

Huff, C. R., & Killias, M. (Eds.). (2013). *Wrongful Convictions and Miscarriages of Justice: Causes and Remedies in North American and European Criminal Justice Systems*. New York, NY: Routledge.

Innocence Project. (2014). *Eyewitness misidentification*. Retrieved from http://www.innocenceproject.org/understand/eyewitness-misidentification.php

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1304-1316.

Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The" general acceptance" of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, *44*(8), 1089-1098.

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the" general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist*, *56*(5), 405-416.

Katzev, R. D., & Wishart, S. S. (1985). The impact of judicial commentary concerning eyewitness identifications on jury decision making. *The Journal of Criminal Law and Criminology*, *76*(3), 733-745.

King, J. F., Zechmeister, E. B., & Shaughnessy, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *The American Journal of Psychology*, 329-343.

Knapp, B. R., Nosofsky, R. M., & Busey, T. A. (2006). Recognizing distinctive faces: A hybrid-similarity exemplar model account. *Memory & Cognition*, *34*(4), 877-889.

Lampinen, J. M., Judges, D. P., Odegard, T. N., & Hamilton, S. (2005). The reactions of mock jurors to the Department of Justice guidelines for the collection and preservation of eyewitness evidence. *Basic and Applied Social Psychology*, *27*(2), 155-162.

Law Reform Commission of Canada, & Brooks, N. (1983). *Police Guidelines, Pretrial Eyewitness Identification Procedures*. Ottawa: Law Reform Commission of Canada.

Leippe, M. R. (1995). The case for expert testimony about eyewitness memory. *Psychology, Public Policy, and Law, 1*(4), 909-959.

Leippe, M. R., Eisenstadt, D., Rauch, S. M., & Seib, H. M. (2004). Timing of eyewitness expert testimony, jurors' need for cognition, and case strength as determinants of trial verdicts. *Journal of Applied Psychology, 89*(3), 524-541.

Levene, H. (1960). Robust tests for equality of variance. In I. Olkin (ed.) *Contributions to Probability and Statistics*. Palo Alto, CA: Stanford.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research, 28*(4), 612-625.

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

Lindsay, R. C. L. (1999). Applying applied research: selling the sequential line-up. *Applied Cognitive Psychology, 13*(3), 219-225.

Lindsay, R. C. L., Lea, J. A., Nosworthy, G. J., Fulford, J. A., Hector, J., LeVan, V., & Seabrook, C. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76*(6), 796-802.

Lindsay, R. C. L., Smith, S. M., & Pryke, S. (1999). Measures of lineup fairness: Do they postdict identification accuracy? *Applied Cognitive Psychology, 13*, S93-S107.

Lindsay, R. C. L. & Wells, G. L. (1980). What price justice? Exploring the relationship between lineup fairness and identification accuracy. *Law and Human Behavior, 4*, 303-314.

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

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

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

Loftus, E. F. (1974). Reconstructing memory: The incredible eyewitness. *Psychology Today*, *8*, 116-119.

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge University Press.

Loftus, E. F., & Ketcham, K. (1991). *Witness for the defence: The accused, the eyewitness, and the expert who put memory on trial*. New York, NY: St. Martin's Press.

Lubbock-Avalanche Journal. (2000, April 16). *Man awaits sentence for wig shop murder*. Retrieved from http://lubbockonline.com/stories/041600/sta_041600053.shtml.

Luus, C. A., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15*(1), 43-57.

Maass, A., Brigham, J. C., & West, S. G. (1985). Testifying on Eyewitness Reliability: Expert Advice Is Not Always Persuasive. *Journal of Applied Social Psychology*, *15*(3), 207-229.

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior*, *5*(4), 299-309.

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

Malpass, R. S., & Lindsay, R. C. (1999). Measuring lineup fairness. *Applied Cognitive Psychology*, *13*, S1-S7.

Malpass, R. S., Ross, S. J., Meissner, C. A. & Marcon, J. L. (2009). The Need for Expert Psychological Testimony on Eyewitness Identification. In B. L. Cutler (Ed.) *Expert Testimony on the Psychology of Eyewitness Identification*. New York, NY: Oxford University Press.

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. E. (2007). Lineup construction and lineup fairness. In R. Lindsay, D. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology*: *Memory for people* (Vol. 2, pp. 155–178). Mahwah, NJ: Erlbaum.

Martire, K. A. (2008). Helping jurors to evaluate eyewitness identifications: The role of expert evidence and judicial instruction. Retrieved from University of New Wouth Wales Library.

Martire, K. A., & Kemp, R. I. (2009). The impact of eyewitness expert evidence and judicial instruction on juror ability to evaluate eyewitness testimony. *Law and Human Behavior*, *33*(3), 225-236.

Martire, K. A. and Kemp, R. I. (2011). Can experts help jurors to evaluate eyewitness evidence? A review of eyewitness expert effects. *Legal and Criminological Psychology*, *16*, 24-36.

McCloskey, M., & Egeth, H. E. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist*, *38*(5), 550-563.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika, 12*(2), 153-157.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3-35.

Moore, E. H. (1948). Elements of error in testimony [Abstract]. *Oregon Law Review, 28*(4), 293-295.

Münsterburg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York, NY: McClure.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691-692.

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

O'Connor, M. C. (2013). *Effects of judicial warnings about cross-race eyewitness testimony on jurors' judgments*. Retrieved from Trinity College Digital Repository, http://ww.digitalrepository.trincoll.edu/theses/342

Olsson, N., Juslin, P., & Winman, A. (1998). Realism of confidence in earwitness versus eyewitness identification. *Journal of Experimental Psychology: Applied, 4*(2), 101-118.

Overbeck, J. L. (2005). Beyond Admissibility: A Practical Look at the Use of Eyewitness Expert Testimony in the Federal Courts. *NYUL Rev., 80*, 1895-1920.

Paterson, H. M., Anderson, D. W. M., & Kemp, R. I. (2013). Cautioning jurors regarding co-witness discussion: the impact of judicial warnings. *Psychology, Crime & Law, 19*(3), 387-304.

Paterson, H. M., Anderson, D. W., & Kemp, R. I. (2013). Cautioning jurors regarding co-witness discussion: the impact of judicial warnings. *Psychology, Crime & Law, 19*(3), 287-304.

Penrod, S., & Cutler, B. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law, 1*(4), 817-845.

Pezdek, K., Avila-Mora, E., & Sperry, K. (2010). Does trial presentation medium matter in jury simulation research? Evaluating the effectiveness of eyewitness expert testimony. *Applied Cognitive Psychology, 24*(5), 673-690.

Pierce, C. A., Block, R. A., & Aguinis, H. (2004). Cautionary note on reporting eta-squared values from multifactor ANOVA designs. *Educational and Psychological Measurement, 64*(6), 916-924.

Ramsey, P. H. (1980). Exact type I error rates for robustness of Student's *t* test with unequal variance. *Journal of Educational Statistics, 5*, 337-350.

Roediger, H. L., Wixted, J. T., & DeSoto, K. A. (2012). The curious complexity between confidence and accuracy in reports from memory. In L. Nadel & W. Sinnott-Armstrong (Eds.), *Memory and law* (pp. 84–118). New York, NY: Oxford.

R v. D.D. (2000), 36 C.R. (5th) 261, 148 C.C.C. (3d) 41, 191 D.L.R. (4th) 60, (2000) 2 S.C.R. 275

R. v. Eakin (1994) OCA No. 901 (QL)

R v. Hay (2013) 365 D.L.R. (4th) 276, 6 C.R. (7th) 215, (2013) 3 S.C.R. 694

R v. Henderson (2009) 67 C.R. (6th) 132 (MBQB)

R v. Henderson (2012) 293 C.C.C. (3d) 479 (MBCA)

R v. McIntosh (1997), 117 C.C.C. (3d) 385 (Ont. C.A.)

R v. Mohan (1994), 89 C.C.C. (3d) 402, 29 C.R. (4th) 243, (1994) 2 S.C.R. 9.

R v. Mohan (1992) 8 O.R. (3d) 173, 55 O.A.C. 309, 71 C.C.C. (3d) 321, 13 C.R. (4th) 292 (Ont. C.A.)

R. v. Nyugen (1997) OCA No. 2801 (QL)

R. v. Stinchcombe (1991) 3 S.C.R. 326

R v. Turner (2012) ONSCJ No. 1634 (QL)

Roediger, H. L., Wixted, J. H., & DeSoto, K. A. (2012). The curious complexity between confidence and accuracy in reports from memory. In L. Nadel & W. Sinnott-Armstrong (Eds.), Memory and law (pp. 84–118). New York, NY: Oxford.

Satterthwaite, F. E. (1946). Of variance components. *Biometrics*, *2*(6), 110-114.

Sauerland, M., & Sporer, S. L. (2009). Fast and confident: postdicting eyewitness identification accuracy in a field study. *Journal of Experimental Psychology: Applied*, *15*(1), 46-62.

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and Type II error properties of the t-test to departures from population normality. *Psychological Bulletin*, *111*(2), 352-360.

Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence*. New York: Doubleday.

Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiential bases of metamemory. *Psychology*, *70*, 567-577.

Smalarz, L. & Wells, G. L. (2013). Post-identification feedback to eyewitnesses impairs

evaluators' abilities to discriminate between accurate and mistaken testimony. *Law and Human Behavior, 38*(2), 194-202.

Smith, E. (1930). The fallibility of eyewitness testimony. *The American Journal of Police Science, 1*(5), 487-495.

Smith, S. M., Lindsay, R. C. L., & Pryke, S. (2000). Postdictors of eyewitness errors: Can false identifications be diagnosed? *Journal of Applied Psychology, 85*(4), 542-550.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*(3), 315-327.

Sporer, S. L. (2001). The cross-race effect: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and Law, 7*(1), 170-200.

State v. Goldberg (95 S.W.3d 345 (Tex. App. - Hous.) [1 Dist.] 2002

Statistics Canada. (2010). *University enrolment by field of study and gender*. Retrieved June 17, 2014 from http://www.statcan.gc.ca/daily-quotidien/100714/t100714a2-eng.htm.

Statistics Canada. (2011). *National household survey*. (Catalogue number 99-004-XWE). Retrieved June 17, 2014 from http://www12.statcan.gc.ca/nhs-enm/2011/dp-pd/prof/index.cfm?Lang=E.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: a meta-analytic comparison. *Law and Human Behavior, 25*(5), 459-473.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates

in police showup and lineup presentations: a meta-analytic comparison. *Law and Human Behavior*, *27*(5), 523-540.

Stinson, V., Devenport, J. L., Cutler, B. L., & Kravitz, D. A. (1997). How effective is the motion-to-suppress safeguard? Judges' perceptions of the suggestiveness and fairness of biased lineup procedures. *Journal of Applied Psychology*, *82*(2), 211-220.

Stuart, D. (2001). *Charter justice in Canadian criminal law.* Toronto: Carswell Publications.

Technical Working Group for Eyewitness Evidence. (1999). Eyewitness evidence: A guide for law enforcement. Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Terrance, C., Thayer, A., & Kehn, A. (2006). Undermining eyewitness confidence inflation: Effecting change through expert testimony. *Journal of Forensic Psychology Practice*, *6*(1), 73-82.

Thompson-Cannino, J. Cotton, R., & Torneo, E. (2010). *Picking Cotton: Our memoir of injustice and redemption*. New York: St. Martin's Press.

Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior*, *22*(2), 217-237.

Tredoux, C. G. (1999). Statistical considerations when determining measures of lineup size and lineup bias. *Applied Cognitive Psychology*, *13*, S9-S26.

Turtle, J. W., Lindsay, R. C. L., & Wells, G. L. (2003). Best practice recommendations for eyewitness evidence procedures: New ideas for the oldest way to solve a case. *The Canadian Journal of Police and Security Services*, *1*(1), 5-18.

Walker, L., & Monahan, J. (1987). Social frameworks: A new use of social science in law.

Virginia Law Review, 73, 559-598.

Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *Journal of Applied Psychology*, *88*(3), 490-499.

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, *36*(12), 1546-1557.

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*(2), 89-103.

Wells, G. L. (1986). Expert psychological testimony: Empirical and conceptual analyses of effects. *Law and Human Behavior*, *10*(1-2), 83-95.

Wells, G. L. (1988). *Eyewitness identification: A system handbook*. Toronto, ON: Carswell Legal Publications.

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, *48*(5), 553-571.

Wells, G. L. (2001). Police lineups: Data, theory, and policy. *Psychology, Public Policy, and Law, 7*(4), 791-801.

Wells, G. L., & Bradfield, A. L. (1999). Measuring the goodness of lineups: Parameter estimation, question effects, and limits to the mock witness paradigm. *Applied Cognitive Psychology*, *13*, S27-S39.

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

Wells, G. L., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88*(3), 776-784.

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

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

Wells, G. L. & Luus, C. E. (1990). Police lineups as experiments: Social methodology as a framework for properly conducted lineups. *Personality and Social Psychology Bulletin*, *16*(1), 106-117.

Wells, G. L. & Murray, D. M. (1984). Eyewitness confidence. In G. L. Wells & E. F. Loftus (Eds). *Eyewitness testimony: Psychological perspectives*, pp. 155-170. New York: Cambridge University Press.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, *54*(1), 277-295.

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

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

Whitley, B. E., & Greenberg, M. S. (1986). The role of eyewitness confidence in juror perceptions of Credibility. *Journal of Applied Social Psychology*, *16*(5), 387-409.

Wise, R. A., Dauphinais, K. A., & Safer, M. A. (2007). A tripartite solution to eyewitness error. *The Journal of Criminal Law and Criminology*, 807-871.

Wogalter, M. S., Malpass, R. S., & Mcquiston, D. E. (2004). A national survey of US police on preparation and conduct of identification lineups. *Psychology, Crime and Law, 10*(1), 69-82.

Yarmey, A. D. (1979). *The psychology of eyewitness testimony*. New York, NY: Free Press.

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

Yarmey, A. D. & Jones, H. T. (1982). Police awareness of the fallibility of eyewitness identification. *Canadian Police College Journal, 6*(3), 113-124.

Zarkadi, T., Wade, K. A., & Stewart, N. (2009). Creating fair lineups for suspects with distinctive features. *Psychological Science, 20*(12), 1448-1453.