ASSESSING THE RELIABILITY OF MULTIPLE-SHOWUP PROCEDURES WITH A SINGLE EYEWITNESS

by

Andrew M. Smith

A thesis submitted to the Department of Psychology

In conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

(August, 2015)

Abstract

In this program of research I examined the influence of multiple-showup procedures on the reliability of eyewitness identification. If an eyewitness rejects a suspect from a showup, law enforcement personnel may search for new suspects, and if a new suspect is found, they may run another showup with the same eyewitness. Law enforcement personnel may go through several iterations of finding suspects and running showups with the same eyewitness for single-perpetrator crimes (Chapter 3, Study 1). The use of multiple-showup procedures with the same eyewitness increased innocence risk (Clark & Godfrey, 2009), the probability that a suspect was innocent given identification (Chapter 3, Studies 2 and 3). The increase in innocence risk was primarily attributable to the fact that innocent suspect identifications cumulate when law enforcement personnel use multiple showups in single-perpetrator crimes. Although pre-showup instructions decreased innocent suspect identifications, innocence risk remained unacceptably high (Chapter 3, Study 4; Chapter 4). On a more positive note, pre-showup instructions intended to reduce eyewitness' perceptions of scarcity (belief that they may not have another opportunity to identify the perpetrator) decreased innocent suspect identifications over and above a may-or-may-not be present admonition (Chapter 4). Scarcity instructions may prove useful in identification procedures, more broadly. I argue that an identification from a multiple-showup procedure is not reasonable evidence of guilt; however, practical constraints may require that law enforcement personnel sometimes show eyewitnesses more than a single showup. A stronger partition between investigative and evidentiary procedures is recommended (Chapter 5).

Co-Authorship

R. C. L. Lindsay

Michelle Bertrand

Natalie Kalmet

Deborah Grossman

Daniel Provenzano

Tiffany Myerson

Acknowledgements

I would like to express my sincere gratitude to my supervisor and mentor Dr. R. C. L. Lindsay. Rod, I cannot thank you enough for everything that you have taught me over the past four years. The many discussions we have had over the past four years helped guide me through the completion of this research, and more importantly helped me to develop as a scientist. To Marilyn Lindsay: thank you for keeping the lab running as smoothly as possible and for reminding Rod of our scheduled meetings. To my committee members, Professors Lisa Dufraimont and Leandre Fabrigar, thank you for providing critical commentary on an earlier draft of this dissertation and for taking the time to serve on my committee. I would also like to thank everyone who assisted in the collection of these data over the past four years, and in particular those who co-authored the manuscripts included in this dissertation: R. C. L. Lindsay, Michelle Bertrand, Natalie Kalmet, Deborah Grossman, Daniel Provenzano, and Tiffany Myerson.

I would also like to thank my family for their encouragement and emotional support throughout this process. And, most importantly, I would like to thank my beloved wife, Kellie (I remembered to use your name this time – see what good training can do!), for her encouragement and emotional support throughout this process. Kellie, thank you for supporting me through the many sleepless nights in which I wrestled with uncertainty.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Innocence advocacy organizations such as the Association in Defense of the Wrongly Convicted (AIDWYC) and the Innocence Project have uncovered hundreds of wrongful convictions in Canada and the United States. At the time of this writing, there are at least 1607 known instances of wrongful conviction in the United States alone (The National Registry of Exonerations, 2015). Mistaken eyewitness identification is a leading antecedent in these miscarriages of justice, having contributed to 33% of them. That eyewitnesses will sometimes identify innocent suspects comes as little surprise to experimental psychologists. Indeed, psychological scientists have long been aware of mistaken eyewitness identification (e.g., Munsterberg, 1908/2009; Whipple, 1909). Over the past several decades, research on eyewitness reliability has increased understanding on an array of issues including the relationship between confidence and accuracy in identification decisions, the moderating effect of feedback on the confidence-accuracy relationship, techniques for interviewing eyewitnesses, the impact of eyewitness expert testimony, and methods for improving the reliability of identification.

In the present dissertation, I extend understanding in the area of identification reliability. Specifically, I examined the police practice of using multiple showup identification procedures with the same eyewitness for single perpetrator crimes. A showup is an identification procedure in which a single person – the suspect – is

presented to an eyewitness for an identification test. If the eyewitness makes an identification, police might make an arrest and begin to investigate the suspect more thoroughly. If the eyewitness does not identify the suspect, it is less clear how police will proceed with the investigation. One possibility is that police will continue to search for the perpetrator and if a new suspect is found, police may use another showup with the same eyewitness. Police may go through several iterations of finding suspects and running showups with the same eyewitness. In point of fact, police might continue this process until they get an identification. It is noteworthy that in one of the most important pieces of American case law as it relates to eyewitness evidence, the eyewitness, Neil, was presented with 30 to 40 photographs of suspects before she identified Biggers as her assailant (*Neil v. Biggers,* 1972). In Chapter 3, I provide empirical evidence that the case of *Neil v. Biggers* (1972) is by no means a one-off or antiquated investigative practice – a survey of North American police officers suggests that this practice is fairly common to this day.

In Chapter 2, I (1) review the existing literature on the reliability of identification, (2) acknowledge that while research on lineups has been extensive over the last several decades, research on showups has lagged behind, (3) review the research that does exist on showup procedures, and (4) provide a brief overview of the research presented in chapters three and four. Chapter 3 is an article my co-authors and I recently published in the *Journal of Experimental Psychology: Applied*. In this article, we present evidence that police officers sometimes use multiple showup procedures with the same eyewitness

and that this practice causes concern for the reliability of identification. In Chapter 4, I present a manuscript my co-authors and I recently submitted to the *Journal of Police and Criminal Psychology*. In this article, we examined how different pre-showup instructions might decrease innocent suspect identifications. Finally, in Chapter 5, I provide general conclusions and suggestions that add to the corpus of recommended best practices for identification procedures.

Chapter 2

Literature Review

Although investigations of eyewitness reliability began over a century ago (e.g., Munsterberg, 1908/2009; Whipple, 1909), ardent investigation of eyewitness reliability did not begin until the 1970s. Classic works by Elizabeth Loftus, (e.g., 1979; Loftus & Palmer, 1974), Robert Buckhout, (1974) and Gary Wells (1978) were particularly influential. Indeed, Wells (1978) made a fruitful distinction between system and estimator variables that has guided research ever since. System variables are those variables over which members of the legal system have some control. For example, the legal system has control over the instructions an eyewitness is given before viewing a lineup. Estimator variables are those variables over which members of the legal system have no control and can therefore only estimate the effects of. For example, the legal system has no control over what time of day a crime occurs. If a crime occurs at night, members of the legal system cannot change this fact, but they can estimate how this might have impacted the eyewitness' ability to successfully encode the perpetrator. Historically, some notable members of the legal system criticized the application of psychological findings to the legal system on the grounds that such findings do not address the abilities of an individual eyewitness in a given instance (e.g., Wigmore, 1909). In appreciation of this criticism, Wells (1978) argued that eyewitness researchers might better assist the criminal justice system through the development of practices that

reduce eyewitness error. Since this time, considerable emphasis has been placed on the development of best practices designed to safeguard against innocent suspect identification. In the next section, I review best-practice recommendations for eyewitness identification.

Best-Practice Recommendations for Eyewitness Identification Procedures

A lineup is an identification procedure in which a suspect and some number of known innocents (hereafter referred to as fillers) are presented to an eyewitness for an identification test. Lineups can be done corporeally, with photographs, or with videos. The purpose of a lineup is to test law enforcement personnel's hypothesis that the suspect perpetrated some crime (Charman & Wells, 2007; Wells & Luus, 1990). Over the past 35 years, researchers have recommended several ways in which law enforcement personnel might modify the traditional lineup in order to decrease innocent suspect identifications whilst (hopefully) maintaining perpetrator identifications. In this section, I discuss several considerations law enforcement must make before constructing a lineup.

Lineup Construction

When constructing lineups, law enforcement personnel must take several factors into consideration. Law enforcement personnel must determine how many people to include in their lineup and how many of these people will be suspects. After determining the lineup size and the number of suspects, law enforcement personnel must determine criteria for selecting fillers. I discuss each of these considerations in turn.

Number of Suspects

Lineups should only include a single suspect. Holding all else constant, the use of multiple-suspect lineups do not increase the risk of mistaken identification for any *given* suspect; however, multiple-suspect lineups do increase the overall risk of mistaken identification, because innocent suspect choices cumulate (Wells & Turtle, 1986). Consider two lineups. Lineup A includes a single suspect and five fillers. Lineup B includes six suspects and no fillers (it is an all-suspect lineup). A random lineup choice from an eyewitness with no memory for the perpetrator will only lead to a suspect identification 16.67% of the time in Lineup A, as fillers siphon 83.33% of guesses away from the suspect (Wells, Smalarz, & Smith, In Preparation; Wells & Turtle, 1986). Because all members of Lineup B are suspects, 100% of choices are suspect choices. All-suspect lineups are like multiple-choice tests in which every response is "correct" (Wells & Turtle, 1986). Even if an all-suspect lineup guaranteed the presence of the perpetrator, a choice from an eyewitness with no memory for the perpetrator would lead to the "identification" of an innocent suspect 83.33% of the time. Including more than one suspect in a lineup greatly increases the likelihood of an innocent suspect identification.

The inclusion of fillers in lineups allows for a more stringent test of law enforcement personnel's hypothesis that some suspect perpetrated some crime. If all members of a lineup are suspects, the hypothesis employed by law enforcement personnel is that the perpetrator is among these lineup members. The scope of the

hypothesis is far too broad as any response other than a lineup rejection supports the hypothesis. In a six-person all-suspect lineup 85.71% (6/7) of response options support law enforcement personnel's hypothesis that the suspect is the culprit. On the contrary, if the lineup includes only a single suspect, the working hypothesis is that this given lineup member (the suspect) perpetrated some crime. The hypothesis is much more narrow in scope as the only response option that supports law enforcement personnel's hypothesis is an identification of the only suspect in the lineup. In a six-person single-suspect lineup, only 14.29% (1/7) of response options support law enforcement personnel's hypothesis. Support for this latter hypothesis is considerably more impressive.

Moreover, lineups should always include fillers because filler identifications are known errors (Wells & Turtle, 1986). Fillers offer a means to impeach incredible eyewitnesses. Indeed, filler identifications are evidence of innocence as they are more likely to occur when the perpetrator is not in the lineup (Wells & Turtle, 1986; Wells, Yang, & Smalarz, 2015). Unless otherwise stated, throughout the remainder of this dissertation I assume that all procedures include only a single suspect.

Lineup Size

After law enforcement personnel determine how many suspects they will include in a lineup, they must determine how many fillers to include. As the size of a lineup increases, so to does the stringency of the hypothesis test. In a standard six-person lineup, the eyewitness has a 16.67% (1/6) chance of selecting the suspect based on chance alone. If one increases the size of a lineup to 24, however, the eyewitness only

has a 4.17% (1/24) chance of selecting the suspect based on chance alone. Generally speaking, suspect identifications obtained from larger lineups are more impressive than those obtained from smaller lineups, because they are less likely to be attributable to chance. The empirical evidence supports this reasoning.

Levi (2007) found that choosing increased from a 12- to 24-person lineup, but because identifications were dispersed over twice as many lineup members, the probability of an innocent suspect identification decreased. Presumably the increased choosing in the 24-person lineup is due to the fact that as the number of response options increases, so too does the probability that at least one of those response options will exceed the eyewitness' decision criterion, the level of match to memory required to elicit a selection. Increasing from 24- to 86- or even 120-person lineups further decreased innocent suspect identifications (Levi, 2007, 2012). Of course, the reduction in innocent suspect identifications did not come without the cost of a loss in perpetrator identifications – increasing lineup size from 12- to 24-persons led to a significant reduction in perpetrator identifications (Levi, 2012). Despite the reduction in perpetrator identifications, identifications from larger lineups were more diagnostic of guilt. That is, the ratio of perpetrator to innocent suspect identifications increased with lineup size. Although an optimal lineup size has not been determined, this is an interesting avenue for future research. Indeed, as lineup size increases, identifications become more reliable; however, there is a need to strike a balance between reliable identifications and not letting too many perpetrators go free.

Filler Selection

Law enforcement personnel must also determine a criterion for selecting fillers. There is little debate that fillers should generally resemble the suspect; however, there is some debate over how law enforcement personnel should achieve this end (e.g., Clark, 2003; Clark, Rush, & Moreland, 2013; Clark & Tunniclif, 2001; Lindsay, Martin, & Webber, 1994; Luus & Wells, 1991; Navon, 1992; Tunniclif & Clark, 2000; Wells, Rydell, & Seelau, 1993). The literature has primarily focused on two methods of filler selection: match-to-suspect and match-to-description-of-the-perpetrator. As the names suggest, match-to-suspect refers to selecting fillers based on their similarity to the suspect. Match-to-description-of-the-perpetrator involves selecting fillers based on their fitting the eyewitness' description of the perpetrator.

The logic behind selecting fillers based on their similarity to the suspect is prefaced on the need to protect innocent suspects. Often law enforcement personnel come to suspect individuals based on their match to an eyewitness' description of the perpetrator. In such instances, innocent suspects would likely share at least some modest resemblance with the eyewitness' memory for the perpetrator. Therefore, if fillers do not bear some general resemblance to the suspect – or the perpetrator – then the innocent suspect might unduly standout as the best match to the eyewitness' memory for the perpetrator. In such instances, an eyewitness might deduce that this is the individual whom law enforcement suspect and infer that this individual is the perpetrator.

There are three potential problems with the match-to-suspect strategy. First, a match-to-suspect strategy may lead to unnecessary similarity among lineup members (Luus & Wells, 1991; Wells et al., 1993). This criterion for filler selection is too stringent. Indeed, it follows from this recommendation that the perfect lineup is one in which the fillers are clones of the suspect (Luus & Wells, 1991). Setting aside hyperbole, the first problem with a match-to-suspect strategy is that adherence to this recommendation might produce lineups in which fillers resemble the perpetrator so closely, that even an eyewitness with a strong memory for the perpetrator cannot discriminate between the perpetrator and the remaining lineup members.

The second problem with a match-to-suspect strategy is the absence of an objective criterion upon which to determine sufficient similarity between the suspect and each lineup member (Luus & Wells, 1991). How similar is similar enough? How do law enforcement personnel determine when the lineup has reached this threshold? A match-to-suspect strategy provides no answer to these questions.

The third problem with a match-to-suspect strategy is that, ironically, it has the potential to lead to lineups in which innocent suspects unduly stand out (Clark & Tunnicliff, 2001; Navon, 1992). Although a filler might appear similar to the suspect, the filler might not possess features described by the eyewitness. To the extent that an eyewitness notices that a filler does not possess a feature that was mentioned in his or her description, the filler might be effectively eliminated from consideration by the eyewitness. Most people would probably take for granted that fillers should share the

same basic demographic characteristics as the suspect (e.g., biological sex, race, age, etcetera). But, on what other features should the fillers resemble the suspect? What if the eyewitness describes – or recalls – that the perpetrator possessed some characteristic not subject to physical measurement?

This is precisely the research question examined by Doob and Kirshenbaum (1973). Specifically, they examined a case in which an eyewitness could describe nothing about two perpetrators except that they were handsome; nevertheless, she was presented with a 12-person lineup and was able to identify the suspect. Even if a match-to-suspect strategy led to the selection of fillers who generally resembled the suspect, there is no guarantee that they will possess the characteristic of handsomeness. This could lead to a lineup in which the suspect is the only member who is handsome. In such an instance, the eyewitness who remembers nothing of the perpetrator other than the fact that he was handsome might pick the only – or most – handsome member in the lineup. Indeed, in the case examined by Doob and Kirshenbaum (1973), participants rated the suspect as the best looking lineup member (Study 1) and in Study 2, he was selected as the most attractive lineup member from a 12-person lineup by 11 of 21 mock witnesses (i.e., eyewitnesses who had no prior exposure to the perpetrator or lineup members but were provided with the description given by the witness).

The match-to-description strategy was intended to overcome these shortcomings (Luus & Wells, 1991; Navon, 1992; Wells et al., 1993). Indeed, the match-to-description strategy has several potential benefits. First, match-to-description should prevent

unwarranted similarity among lineup members as it suggests that fillers should only share with the suspect, those features that were described by the eyewitness. The idea is to produce a lineup in which fillers have only limited similarity to the suspect. Lineups with such "propitious heterogeneity" should lead to more perpetrator identifications than lineups produced from a match-to-suspect strategy, because they better facilitate the detection of diagnostic features (Wells et al., 1993). Yet, description-matching innocent suspects should not be at increased risk of mistaken identification because all lineup members should match descriptions of perpetrators. Second, and related to the first point, a match-to-description strategy provides law enforcement personnel with an objective criterion to determine sufficient similarity among lineup members (Luus & Wells, 1991; Wells et al., 1993). Namely, fillers should resemble the perpetrator on those features – and only those features – described by the eyewitness. Third, if all lineup members fit the eyewitness' description of the perpetrator, this should prevent the eyewitness from being able to discard fillers on the grounds that they are missing features described by the eyewitness (Navon, 1992).

Although the match-to-description strategy sounds promising in theory, empirical results on the efficacy of this procedure are mixed. Wells et al. (1993) found that description-matched lineups increased perpetrator identifications without increasing innocent suspect identifications. Yet, subsequent research provides scant support for the superiority of description-matched lineups. Clark et al. (2013) examined six studies (and seven comparisons) comparing description- and suspect-matched lineups. Five of the

seven hypothesis tests showed evidence of increased probative value in suspect-matched lineups.[1] Meta-analytically, the description-matched lineups led to an increase in both perpetrator and innocent suspect identifications, but innocent suspect identifications increased to a proportionally greater extent leading to the reduction in probative value (Clark et al., 2013). These results favor suspect-matched lineups, yet, it is also clear from the findings of Doob and Kirshenbaum (1973) that innocent suspects are in danger of mistaken identification if fillers do not match the eyewitness' description of the perpetrator. Taking this into consideration, the ideal procedure probably involves an amalgam of matching fillers both to the suspect and to the eyewitness' description (Clark et al., 2013). Yet, there is a dearth of research on the topic of filler selection and there is a dire need to determine the most appropriate methods of filler selection and to examine potential moderators.

Presentation Methods

Two methods of lineup presentation have dominated the eyewitness literature: simultaneous and sequential presentation. The simultaneous lineup is what most lay people would envision if one asked what a police lineup looked like. The simultaneous lineup involves presenting lineup members to the eyewitness in an array. Hence, the eyewitness is presented with all lineup members "simultaneously" and eyewitnesses

[1] By probative value, I am referring to a collection of related measures derived from Bayes' Theorem including diagnosticity (Wells & Lindsay, 1980), proportion guilty (Pryke, Dysart, Dupuis, & Lindsay, 2004), and innocence risk (Clark & Godfrey, 2009). Diagnosticity is the *likelihood* that a suspect is guilty given identification, proportion guilty is the *probability* that a suspect is guilty given identification, and innocence risk is the *probability* that a suspect is innocent given identification (or 1 – proportion guilty).

might compare and contrast lineup members to determine the best match to memory (i.e., make relative judgments, Wells, 1984). If the best match to memory is above some internal decision criterion, the eyewitness will identify this lineup member (Clark, 2003). While this practice is (arguably) not problematic when the perpetrator is present – because the perpetrator will generally be the best match to memory – it is very problematic when the perpetrator is absent from the lineup. At the extreme, if eyewitnesses *always* selected the best match to memory (i.e., the eyewitness' response criterion was infinitely lax), this would lead to the identification of many innocent suspects.

In an effort to discourage eyewitnesses from using relative judgments, Lindsay and Wells (1985) devised the sequential lineup procedure. The sequential lineup has four key components: (1) lineup members are presented to the eyewitness one at a time, (2) the eyewitness must make a decision about each lineup member before proceeding to the next, (3) eyewitnesses cannot return to previously rejected lineup members, and (4) the eyewitness does not know how many members the lineup includes (i.e., the lineup is backloaded). The first three components of the sequential lineup were intended to prevent intra-lineup comparisons (i.e., relative judgments) and the fourth component, backloading, was intended to prevent eyewitnesses who had not identified a lineup member from becoming more willing to choose as the end of the lineup approached (i.e., from liberally shifting their decision criteria).

No topic in the eyewitness identification literature is more debated than the mode in which lineup members should be presented to the eyewitness (e.g., Clark, 2012; Gronlund, Wixted, & Mickes, 2014; Lindsay, Mansour, Beaudry, Leach, & Bertrand, 2009; Malpass, Tredoux, & McQuiston-Surrett, 2009; Wells, 2014; Wells, Steblay, & Dysart, 2012; Wixted & Mickes, 2014). Early comparisons of simultaneous and sequential lineups found overwhelming support for the sequential lineup procedure. The sequential lineup reduced innocent suspect identifications – presumably by decreasing the tendency of eyewitnesses to use relative judgments – and increased the diagnostic value of identifications (Cutler & Penrod, 1988; Lindsay & Wells, 1985). The sequential lineup also reduced perpetrator identifications; however, the decrease in perpetrator identifications was proportionally less than the decrease in innocent suspect identifications (hence the increased diagnostic value of the sequential lineup; Cutler & Penrod, 1988; Lindsay & Wells, 1985).

This early research on the efficacy of sequential presentation is supported by two meta-analyses on the topic that found the same pattern of results (Steblay, Dysart, Fulero, & Lindsay, 2001; Steblay, Dysart, & Wells, 2011), but others have challenged this pattern of results. Clark (2012) meta-analyzed a different subset of those studies examined by Steblay et al. (2011) and five additional studies and reached a different conclusion: the sequential lineup leads to proportionally equivalent decrease in both innocent and correct identifications and therefore does not increase the diagnosticity of an identification decision.

Palmer and Brewer (2012) also meta-analytically compared simultaneous and sequential lineups, but within the architecture of signal detection theory. Before considering Palmer and Brewer's meta-analysis, a brief discussion of signal detection theory (SDT e.g., Green & Swets, 1966; MacMillan & Creelman, 2005) is in order. SDT permits the dissociation of two phenomena, discriminability and response bias. Discriminability is a measure of one's ability to differentiate between the presence and absence of some signal in the presence of noise. Response bias is the tendency for one to favor one response over some other response independent of ability to discriminate. In an identification task, a procedure promotes discriminability to the extent that it leads eyewitnesses to identify the perpetrator when present and reject the lineup when the perpetrator is absent. The response bias associated with an identification procedure is a measure of the extent to which it promotes choosing (the suspect or fillers). A lineup procedure has a conservative response bias to the extent that it promotes lineup rejections and has a liberal (or lax) response bias to the extent that it promotes identifications of lineup members.

Traditional SDT measures (e.g., ROC analysis, calculation of *d'* and *c* for detection tasks) are not appropriate for lineup tasks. Lineup procedures entail a compound-decision task (Duncan, 2006; Palmer & Brewer, 2012). Eyewitnesses must complete both a detection task in which they decide whether the perpetrator is present and if an eyewitness concludes that the perpetrator is present, the eyewitness must then complete an identification task by determining which of the lineup members is the

perpetrator (Duncan, 2006; Palmer & Brewer, 2012).[2] As I will discuss shortly, sometimes researchers ignore the fact that lineups are a compound decision task and treat filler identifications as synonymous with rejections for computational simplicity. The problem with this practice – beyond the fact that it is nonsensical to treat false positives as correct rejections – is that treating filler identifications as correct rejections leads to underestimations of the leniency of response bias (i.e., it underestimates how willing eyewitnesses were to choose); and thus, it also leads to overestimations of discriminability (Wells et al., In Preparation). Compound-SDT overcomes these limitations as all affirmative responses are included in estimates of discriminability and response bias. Using this procedure, Palmer and Brewer's (2012) meta-analysis indicated that the sequential advantage was due to a stringent shift in response bias. While both simultaneous and sequential lineups exacted lenient response biases, sequential lineups were less lenient and better reflected the target-present base rate (50%), which led to more accurate identification decisions (Palmer & Brewer, 2012).

Over the past few years, several researchers have compared simultaneous and sequential lineups using receiver operating characteristic (ROC) analysis (Andersen, Carlson, Carlson, & Gronlund, 2014; Carlson & Carlson, 2014; Dobolyi & Dodson, 2013; Gronlund et al., 2012; Mickes, Flowe, & Wixted, 2012). ROC is based on SDT and is said to be a nonparametric measure of discriminability (Egan, 1958). In eyewitness tasks, ROC curves are generated by plotting a series of points in two-

[2] A thorough review of compound signal detection tasks is beyond the scope of the current dissertation, but interested readers are referred to Duncan, 2006; MacMillan & Creelaman, 2005; Palmer & Brewer, 2012.

dimensional space with ordinate values representing the proportion of perpetrator identifications and abscissa values representing the proportion of innocent suspect identifications associated with some level of confidence. The left-most point in the ROC space represents the proportion of perpetrator and innocent suspect identifications associated with the highest level of confidence. Working from left to right, the next point in the ROC space represents the proportion of perpetrator and innocent suspect identifications associated with the two highest levels of confidence. One continues plotting points in this manner until the right-most point in the ROC space represents the proportion of perpetrator and innocent suspect identifications collapsed over all levels of confidence. Thus, as one moves from left to right in the ROC space, points are associated with more lenient decision criteria. After plotting points in the ROC space, an ROC curve is fit to these points. When comparing ROC curves, the curve that bows closer to the upper left corner of the ROC space is said to be associated with superior discriminability. That is, the procedure with a greater area under the curve (AUC) presumably produces superior discriminability.

All ROC comparisons of simultaneous and sequential lineups that I am aware of show some evidence of a *simultaneous* superiority effect (Andersen et al., 2014; Carlson & Carlson, 2014; Dobolyi & Dodson, 2013; Gronlund et al., 2012; Mickes et al., 2012). Why do studies employing ROC analysis consistently show a pattern of results directly opposite to what has been found in two meta-analyses (Steblay et al., 2001; Steblay et al., 2011)? Implicitly, proponents of ROC analysis suggest that the divergent results are

attributable to the fact that ROC analysis is superior to the Bayesian measures (e.g., proportion guilty, Pryke et al., 2004; diagnosticity, Wells & Lindsay, 1980) traditionally used in eyewitness research (e.g., Gronlund, Wixted, & Mickes, 2014; Wixted & Mickes, 2012). The strongest support for the claim that the discrepancy in results is attributable to a difference in the quality of measures would come from a demonstration that the same dataset shows an advantage for one procedure (e.g., the sequential lineup) when using Bayesian measures and an advantage for the other procedure when using ROC analysis (e.g., the simultaneous lineup). None of the ROC comparisons of simultaneous and sequential lineups have demonstrated this phenomenon. In point of fact, each study using ROC analysis to compare simultaneous and sequential lineups has found that simultaneous lineups lead to equivalent (e.g., Dobolyi & Dodson, 2013; Gronlund et al., 2012; Mickes et al., 2012 Experiment 2) or even superior (Andersen et al., 2014; Carlson & Carlson, 2014; Mickes et al., 2012 Experiments 1A and 1B) diagnosticity when compared to sequential lineups.

It is not unheard of for simultaneous lineups to produce equivalent or even superior diagnosticity when compared to sequential lineups, however, the pattern of results is relatively rare (Steblay, 2001; Steblay, 2011). Accordingly, there is scant evidence that ROC comparisons have uncovered some latent effect that has gone unnoticed in the literature. Rather, the evidence suggests that the ROC advocates have failed to replicate the traditional pattern of results. Future research should examine potential moderators of the sequential- and simultaneous-superiority effects.

Although the sequential lineup advantage has recently been challenged, the bulk of the literature does suggest that sequential lineups decrease innocent suspect identifications and increase the diagnosticity of identification decisions (Steblay et al., 2001; Steblay et al., 2011). Accordingly, it seems prudent, that, without further evidence of simultaneous superiority and without knowing how the tacit assumptions of ROC analysis in the context of lineup procedures influence conclusions, eyewitness researchers should not revise recommendations that law enforcement personnel present lineup members sequentially.

The May-or-May-Not Be Present Admonition

If an eyewitness immediately recognizes someone in a lineup, the eyewitness is likely to identify that individual; but what will an eyewitness do if he or she does not immediately recognize anyone in the lineup? One possibility is that the eyewitness will search for reasons why none of the lineup members stand out as the perpetrator (e.g., change in appearance, length of delay, poor encoding conditions, etc.). To the extent that the eyewitness believes that the perpetrator is present, the eyewitness might resort to comparing and contrasting lineup members with one another and selecting the lineup member who most closely resembles his or her memory for the perpetrator (i.e., a relative-judgment strategy, Wells, 1984). As articulated above, this might not be problematic when the perpetrator is present, but will increase the incidence of innocent suspect identification when the target is absent. If the eyewitness is instructed before the identification procedure that the perpetrator may-or-may-not be present, however, then

the eyewitness now has another potential explanation for why he or she does not recognize any of the lineup members – because the perpetrator is not present (Charman & Wells, 2007; Steblay, 2013).

Comparisons of biased instructions (instructions which imply the presence of the perpetrator) and unbiased instructions (instructions which do not imply the presence of the perpetrator) have received considerable attention as is evidenced by four meta-analyses on the topic (Clark, 2005, 2012; Steblay, 1997, 2013). Over time, however, the key manipulation has become the presence or absence of the may-or-may-not be present admonition (Steblay, 2013). Cautioning eyewitnesses that the perpetrator may-or-may-not be present, decreases innocent suspect identifications and increases the diagnosticity of identification decisions (Steblay, 2013). Yet, it is important to acknowledge that the may-or-may-not admonition is not a panacea as this does not abolish innocent suspect identifications and its utility may be diminished in the presence of pre-admonition suggestion (Quinlivan, Neuschatz, Cutler, Wells, McClung, & Harker, 2012).

In addition, the may-or-may-not admonition does lead to a reduction in perpetrator identifications (Clark, 2012); however, the loss in correct identifications is not as problematic as it might appear at face value, because many of these identifications are due to strategic guessing (see Wells et al., 2012). Consider the nine perpetrator-present comparisons reported by Steblay (2013) in which the authors report the proportion of decisions leading to a filler identification. In the presence of the may-or-may-not admonishment, 69% of eyewitnesses identified a lineup member – 54% correctly chose

the perpetrator and 15% incorrectly chose a filler. In the absence of the may-or-may-not admonition, 85% of eyewitness identified a lineup member – 59% correctly chose the perpetrator and 26% identified a filler. Thus, although eyewitnesses who were not admonished chose (85%) 16% more frequently from target-present lineups than those eyewitnesses who were admonished (69%), the difference in perpetrator identifications (5%) was considerably less. Only about one in every three additional identifications that resulted from not admonishing eyewitnesses were of the perpetrator, reflecting the fact that these are not reliable identification decisions – they are probably better classified as strategic guesses.

Single- vs. Double-Blind Administration

The long history of experimenter expectancy effects in psychological research (Rosenthal & Rubin, 1978) suggests that lineup administrators should not know the identity of the suspect in a lineup procedure (Austin, Zimmerman, Rhead, & Kovera, 2013). Experimenters may inadvertently convey their hypotheses to participants and knowledge of the experimenter's hypothesis might influence the decision of the eyewitness. At the extreme, if a lineup administrator knew the identity of the suspect and told the eyewitness whom she suspected, the lineup would functionally become a showup. If the eyewitness is informed that the police suspect a given lineup member, there is little reason for the eyewitness to pay attention to the remaining lineup members.

To date, only one study has compared single- and double-blind lineup administration in both target-present and –absent lineups (Greathouse & Kovera, 2009).

Greathouse and Kovera (2009) manipulated target presence, administrator knowledge (single- v. double-blind), presentation method (simultaneous v. sequential), and the pre-identification admonishment. The authors found a three-way interaction between administrator knowledge, presentation method, and the pre-identification admonishment. When the lineup conditions were biased (used simultaneous presentation and did not provide an admonishment), administrator knowledge increased suspect choices (both perpetrator and innocent suspects); however, this was not a result of administrator knowledge increasing the tendency for an eyewitness to choose. Rather, when administrators knew the identity of the suspect, they "steered" eyewitnesses away from fillers and towards the suspect. Indeed, when the administrator did not know the identity of the suspect, 32% of eyewitnesses identified the suspect and 58% identified a filler. When the administrator knew the identity of the suspect, 60% of eyewitnesses identified the suspect and only 30% identified a filler. Moreover, identifications were twice as diagnostic when administrators did not know the identity of the suspect (6.66) compared to when administrators knew the identity of the suspect (3.25, Greathouse & Kovera, 2009).

This is a powerful demonstration of why administrators should be blind to the identity of the perpetrator. Administrator knowledge will simply lead to the identification of far too many innocent suspects. Furthermore, the purpose of a lineup is to test the hypothesis that the suspect perpetrated some crime. Increasing the suggestiveness of an identification procedure confounds the experiment, leaving the triers

of fact with the impossible task of trying to sort out how much of the identification is attributable to the fact that the suspect is the perpetrator and how much of the identification is attributable to the suggestiveness of the procedure (Smith & Dufraimont, 2014). As an aside, it is also notable that administrator knowledge did not produce the same adverse effects when lineup members were presented sequentially and eyewitnesses were admonished (Greathouse & Kovera, 2009).

Showups

The literature reviewed to this point highlights an array of factors eyewitness researchers have examined over the past 40 years. Eyewitness researchers have made considerable progress in developing best-practice recommendations for lineups. There is ample evidence that sequential presentation, the may-or-may-not be present admonition, and double-blind presentation all decrease innocent suspect identifications and increase the reliability of identification. Those who have contributed to this research should be commended for their efforts. Yet, lineups are not the only medium for an identification test. Indeed, law enforcement personnel often use showups to test their hypothesis that some suspect perpetrated some crime. Despite their common use, showups have not received sufficient empirical attention.

While research on lineup identification procedures has exploded over the last 40 years, showups have received comparatively little empirical attention. Indeed, a PsycINFO search for the terms “lineup” or “line-up” and “eyewitness” returned 625 results. By comparison, an analogous PsycINFO search for the terms “showup” or

"show-up" and "eyewitness" returned 26 results. Worse yet, only 19 of these results were empirical journal articles. This asymmetry in research focus is in spite of the fact that the extant literature would suggest that both lineups and showups are common (e.g., Davis, Valentine, Memon, & Roberts, 2015; Gonzlez, Ellsworth, & Pembroke, 1993). The majority of research on showup procedures has focused on comparing them with lineups (Beal, Schmitt, & Dekle, 1995; Dekle, Beal, Elliott, & Hunneycutt, 1996; Gonzalez et al., 1993; Lindsay, Pozzulo, Craig, Lee, & Corber, 1997; Wagenaar & Veefkind, 1992; Wetmore et al., 2015; Yarmey, Yarmey, & Yarmey, 1994, 1996), on examining their impact on later lineup identification procedures (Behrman & Vayder, 1994; Haw, Dickinson, & Meissner, 2007; Valentine, Davis, Memon, & Roberts, 2012), and on the influence of clothing bias (Dysart, Lindsay, & Dupuis, 2006; Lawson & Dysart, 2014; Wetmore et al., 2015; Yarmey et al., 1996) and other contextual influences (Smith, Leach, & Cutler, 2013). I now briefly review research comparing lineups and showups before proceeding to examine the influence of contextual variables on identification performance in showups.

Showups vs. Lineups

The legal system has long held that showup procedures are inherently suggestive (e.g., *Brathwaite v. Manson,* 1977; *Neil v. Biggers,* 1972; *Stovall v. Denno,* 1967). Showups are inherently suggestive because only a single individual, the suspect, is presented to the eyewitness. Accordingly, it is readily apparent to the eyewitness whom police suspect committed the crime. Moreover, because only a single individual is

presented to the eyewitness, the showup procedure also offers a weaker test of law enforcement personnel's hypothesis that the suspect is the perpetrator than do lineups and afford no protection to potentially innocent suspects. Consider an eyewitness with no memory for the perpetrator who is nevertheless willing to guess the identity of the perpetrator. In a six-person lineup, the eyewitness will "identify" the suspect 16.67% (1/6) of the time. In a showup, the eyewitness will "identify" the suspect 100% (1/1) of the time. Otherwise put, if an eyewitness with no memory for the perpetrator were willing to choose a lineup member, 83.33% (5/6) of decisions would impeach the credibility of the eyewitness and not further endanger a potentially innocent suspect. By contrast, the showup has no means by which to impeach the credibility of the eyewitness who is willing to guess or deduce that the suspect is the perpetrator.

The empirical literature by and large converges with the logic that showups are inferior to lineup procedures. Steblay et al. (2003) meta-analytically compared lineups and showups and found that the perpetrator was equally likely to be identified in both procedures, but showups increased the risk of innocent suspect identifications even though false positives were more common in lineups. Although false positives were more common in lineups, lineups still afforded innocent suspects greater protection because many false positives were filler choices. On the contrary, in a showup procedure all false positives are innocent suspect choices.

More recently, Wetmore et al. (2015) questioned whether there might be instances in which showups are superior to lineups. Specifically, Wetmore et al. (2015) examined

two potential moderators: (1) the length of delay between a crime and an identification procedure, and (2) the extent to which a lineup is biased. Showups can be (and often are) conducted very shortly after a crime occurs, whereas lineups generally require a longer delay (e.g., Behrman & Davey, 2001; Davis et al., 2015). Wetmore et al. (2015) questioned whether showups conducted after a short delay might be superior to lineups conducted after a long delay. Lineups were superior to showups even in the instance that showups were conducted immediately after encoding and lineups were conducted two days after encoding (Wetmore et al., 2015).

A lineup is biased to the extent that the number of lineup members resembling the perpetrator (the functional size) is less than the total number of lineup members (the nominal size, Wells, Leippe, & Ostrom, 1979). Because showups generally result in less choosing than lineups (Steblay et al., 2003), Wetmore et al. (2015) speculated that showups might be advantageous to biased lineups. Yet, both fair and biased lineups proved to be superior to showups (Wetmore et al., 2015).

Given that the empirical literature suggests that lineups are uniformly superior to showups, some (e.g., Agricola, 2009; Wetmore et al., 2015) have argued that showups might be prohibited. Certainly, when it is practical to use a lineup, law enforcement personnel should not use a showup. But it is not always practical to use a lineup. For example, what are law enforcement personnel to do when they locate an individual fitting a general description of the perpetrator near the scene of a crime in both time and space? Match-to-a-general description is not probable cause for arrest (*Washington v. Lambert,*

1996; *R. v. Mann*, 2004), but it is cause for investigatory detention. When a suspect is placed under investigatory detention, law enforcement personnel have a sufficient amount of time to run a showup procedure and if the eyewitness identifies the suspect, this establishes probable cause for arrest. Investigatory detention does not provide law enforcement personnel with a sufficient amount of time to construct and run a proper lineup (Wells & Wilford, 2013). Accordingly, showups appear to fill an important function in the criminal justice system.[3]

Contextual Influences in Showups

Given that showups have a role in the criminal justice system, it is important to consider best-practice recommendations that safeguard against innocent suspect identifications. While some recommended best-practices for lineup procedures are likely transferable to showup procedures (e.g., use of a pre-identification admonition), showups also come with a host of unique problems. Given that showups are generally used in the field (Davis et al., 2015; Behrman & Davey, 2001), law enforcement personnel have less control over their environments and contextual variables that may be controlled for in lineup procedures might influence decisions in showups. One such variable is the clothing worn by the suspect. It is very common for eyewitnesses providing descriptions of perpetrators to describe the clothing the perpetrator was wearing (Lindsay et al., 1994, Study 1). Locating someone near the scene of the crime wearing clothing similar to the

[3] There are a countless number of other exigent circumstances in which showups might have utility; however, the above scenario seems to be the most common.

eyewitness' description of the perpetrator's clothing may be cause for a showup procedure.

Presenting a suspect to the eyewitness in the same (or highly similar) clothing to that worn by the perpetrator increases innocent suspect identifications (Dysart et al., 2006; Wetmore et al., 2015; Yarmey et al., 1996); though there is some evidence that this only occurs when clothing is distinct (Dysart et al., 2006). The influence of clothing bias on perpetrator identifications is less clear. While some authors have found no effect of clothing bias on perpetrator identifications (Dysart et al., 2006; Yarmey et al., 1996), others have found that it increases perpetrator identifications (Lawson & Dysart, 2012; Wetmore et al., 2015). Indeed, Wetmore et al. (2015) argued that presenting suspects in the same clothing that was worn by the perpetrator increased discriminability. Likewise, other recent research indicated that presenting property stolen at encoding alongside the suspect increased discriminability (Smith et al., 2013).

One (of many) possible explanation for the divergent effects of clothing bias on perpetrator identifications is based on how encoding conditions may have differed across studies. To the extent that conditions are favorable to encoding the perpetrator (e.g., clear view, good lighting, sufficient length of exposure, etc.), clothing bias might not increase perpetrator identifications because the perpetrator will "outshine" the clothing cue (Smith, 1988, 1994). To the extent that encoding conditions are less favorable, clothing may aid eyewitnesses in identifying the perpetrator. This might be an avenue worth exploring in future research. Despite the lack of convergent findings in reference to

perpetrator identifications, the research does consistently demonstrate that clothing bias increases innocent suspect identifications. Accordingly, law enforcement personnel should make efforts to avoid presenting suspects in clothing that is similar to the description of the perpetrator's clothing.

The Present Research

In the present dissertation, I added to the agglomeration of knowledge on showup procedures and more generally on the use of multiple-identification procedures. Past research has shown that presenting a suspect to an eyewitness in a showup (e.g., Behrman & Vayder, 1994; Haw et al., 2007) or mugshot (Dysart, Lindsay, Hammond, & Dupuis, 2001; Goodsell, Neuschatz, & Gronlund, 2009) increases the probability that the eyewitness will identify that suspect if presented in a later lineup. Yet, there is no research to my knowledge that has examined the impact of presenting eyewitnesses with multiple identification procedures each including a novel suspect. This is despite the fact that over the course of an investigation, law enforcement personnel might come to suspect multiple persons (e.g., *Neil v. Biggers*, 1972).

In the present dissertation, I examined how the use of multiple-showup procedures impacts the reliability of identification decisions. In the first of four studies included in Chapter 3, I established that, if an eyewitness rejects a suspect from a showup, law enforcement personnel might find another suspect and run another showup with the same eyewitness. Law enforcement personnel may go through several iterations of finding suspects and running showups with the same eyewitness for single-perpetrator

crimes. In studies two and three I examined how this practice impacts the reliability of identification decisions and in study four, I examined if a pre-showup admonition could improve the reliability of identification decisions when multiple showups are employed with the same eyewitness.

In Chapter 4, I present a study in which I modified the pre-identification instructions in an effort to further decrease innocent suspect identifications. Further, I examined whether disabusing eyewitnesses of the notion that opportunities to identify the perpetrator are scarce could decrease innocent suspect identifications over and above a may-or-may-not admonition. Finally, in Chapter 5, I provide concluding remarks and practical recommendations on the use of multiple-showup procedures.

Chapter 3

The Impact of Multiple Showups on Eyewitness Decision-Making and Innocence Risk[4]

After crimes are committed, eyewitnesses routinely provide police with descriptions of events and perpetrators. Police may search the vicinity of a crime for individuals who fit the description of the perpetrator. If police locate such an individual, they may conduct a showup. A showup is an identification procedure in which police present a single individual, the suspect, to an eyewitness and ask the eyewitness if the suspect is the perpetrator. Showups are inherently suggestive. Unlike lineups in which a single suspect is embedded amongst fillers (known innocents), showups include only a single individual (the suspect). Thus, the identity of the suspect is evident to the eyewitness. Moreover, research comparing showups and lineups indicates that showups put innocent suspects at greater risk of mistaken identification (Steblay, Dysart, Fulero, & Lindsay, 2003) and increase the probability that a suspect is innocent given identification (referred to as "innocence risk"; Clark & Godfrey, 2009).

Although the increased risk of innocent suspect identifications might lead some to conclude that showups ought to be disavowed (e.g., Agricola, 2009), this fails to consider practical constraints. Particularly, including a suspect in a lineup can require greater evidence of guilt than that required for a showup. In order to be legally detained long enough to conduct a live lineup, suspects need to be placed under arrest (Wells & Wilford, 2013) but locating an individual who fits a general description of the perpetrator near the scene of the crime is not probable cause for arrest (*Grant* v. *Long Beach*, 2002). Even if photos of the individual already exist in police records, creating a photo array and presenting it to the eyewitness is time consuming. Showups, however, can be conducted very quickly and do not require that the suspect be detained for an extensive period of time. Thus, showups offer the potential to promptly exculpate the innocent and swiftly inculpate the guilty. Given that showups fill an investigative function not easily accomplished with lineups, there is good reason to examine variables that influence decision-making in showups, as opposed to solely comparing their costs and benefits relative to lineups.

Indeed, some research has examined contextual variables that affect the quality of eyewitness decision-making in showups. For example, presenting a suspect in highly similar clothing to that worn by the perpetrator increases innocent suspect identifications (Yarmey, Yarmey, & Yarmey, 1996) but only when the clothing is distinctive (Dysart, Lindsay, & Dupuis, 2006); however, biased clothing does not increase perpetrator identifications (Dysart et al., 2006; Yarmey et al., 1996). Though some research has

examined the influence of contextual variables on the rates of perpetrator and innocent suspect identifications (i.e., hits and false alarms), there is a dearth of research examining how police practices affect these rates, the interest of the present manuscript.

When showups are conducted shortly after the crime, police have likely invested effort in questioning the eyewitness and examining the crime scene, but have likely expended little effort in investigating a single suspect. Often the suspect is included in a showup opportunistically, due to fitting the eyewitness' description of the perpetrator and being near the scene of the crime. When an eyewitness identifies a suspect from a showup, police will arrest the suspect and continue their investigation. But, what do police do when an eyewitness rejects a suspect in a showup? One possibility is that police might find a new suspect who fits the description of the perpetrator and present this suspect in a showup. Likewise, if the eyewitness rejects the second suspect, police may recruit a third suspect and conduct another showup. Police may continue this iterative process until the eyewitness identifies a suspect, even over the span of months.

Indeed, in *Neil v. Biggers* (1972), the victim was presented with several lineups and showups including 30 to 40 photographs of suspects over a seven-month period before identifying Biggers as her assailant. This confirms our speculation that if an eyewitness rejects a suspect from an identification procedure, police may find a new suspect to present to the eyewitness. Moreover, we surveyed North American police officers (Study 1) and a sizable minority who had used showups in the past year reported having presented the same eyewitness with multiple showups for a single perpetrator

crime at least once within the same calendar year. This leads us to question how the use of multiple showups impacts the probability of innocent suspect and perpetrator identifications.

Single- vs. All-Suspect Models

Wells and Turtle (1986) distinguished between *single*- and *all*-suspect lineups. The single-suspect lineup is a model in which a single suspect is embedded amongst fillers. The all-suspect lineup is a model that does not include fillers – each lineup member is a suspect. The risk of mistaken identification for a *given* suspect differs little for single- and all-suspect models; however, because all-suspect lineups are void of fillers any lineup choice is a suspect identification, and thus, the overall probability of a false identification (i.e., "lineup-wise error") increases (Wells & Turtle, 1986). Over time, all-suspect lineups will lead to many more innocent suspect identifications than will single-suspect lineups.

Conducting multiple showups with the same eyewitness for a single-perpetrator crime might produce analogous problems to that of an all-suspect lineup. By definition, a showup is an all-suspect model. What differentiates a showup from an all-suspect lineup is that a showup contains only a single individual, whereas lineups contain multiple

individuals.[5] Providing eyewitnesses who reject suspects from showups with additional suspects will necessarily increase the overall probability of a false identification. This is a mathematical certainty so long as the probability of a false identification on any given showup exceeds zero.

Moreover, as the number of innocents presented before the perpetrator increases, the probability of a perpetrator identification will decrease. If false identifications increase as the number of showups increase, then the number of eyewitnesses that will survive to see the perpetrator will decrease. Therefore, the number of eyewitnesses who will identify the perpetrator will also decrease. This problem will be exacerbated the deeper the perpetrator is in a queue of potential showups.

Despite the obvious downside of conducting multiple showups, exigent circumstances may, at times, warrant their use. Accordingly, it is not sufficient to establish that innocent suspect identifications increase and that perpetrator identifications decrease. It is also necessary to examine the extent to which innocent suspect identifications increase and perpetrator identifications decrease. If a procedure is to be deemed sensible, then it is necessary to demonstrate that guilty suspects are more likely to be identified from that procedure than are innocent suspects (i.e., innocence risk < .50).

---

[5] We do not want to dwell on the differences between showups and lineups as that is not the topic of the present paper, but there are important differences between multiple showups and all-suspect sequential lineups. First, when presented with the initial showup, it is unlikely that eyewitnesses would expect to have additional opportunities to identify the perpetrator. Although an eyewitness in a sequential lineup does not know how many members will be presented in that lineup, the eyewitness does know that a series of persons will be presented. Accordingly, it is quite probable that eyewitnesses would use different decision strategies to respond to the different tasks. Second, each showup is its own procedure. Unlike a sequential lineup in which each lineup member is presented to the eyewitness immediately upon rejection of the prior, there will always be a delay between showups while police search for additional suspects. The delay may be as short as a few minutes, but it could extend over months as in the case of *Neil v. Biggers* (1972).

To this end, we developed a cumulative measure of innocence risk (Clark & Godfrey, 2009) to assess the sensibility of conducting multiple showups with the same eyewitness for single perpetrator crimes. Our cumulative measure of innocence risk reflects that not all eyewitnesses will see the perpetrator if the perpetrator is not presented in the initial showup, and that innocent suspect identifications from multiple showups are additive. We discuss our cumulative measure of innocence risk further in Study 2.

Attrition of Low Criterion-Setting Eyewitnesses

Although the probability of a false identification will increase with the use of multiple showups, research examining the utility of blank lineups (Palmer, Brewer, & Weber, 2012; Wells, 1984) suggests that a given suspect presented after the initial showup might be at less risk of identification than the suspect presented in the initial showup. A blank lineup contains only fillers and is presented before a suspect-containing lineup. Those who rejected the blank lineup were less likely to choose from the suspect-containing lineup than were those who were not presented with a blank lineup (Palmer et al., 2012, Studies 1a and 1b; Wells, 1984).

The blank lineup reveals those eyewitnesses who are relatively more willing to choose (low criterion-setting) and those who are relatively less willing to choose (high criterion-setting). Therefore, it is not surprising that those who rejected the blank lineup chose less frequently from the suspect-containing lineup than did those eyewitnesses who were not exposed to a blank lineup. By comparing only those who rejected the blank lineup to those who were not exposed to a blank lineup, one is comparing a filtered group

of high criterion-setting eyewitnesses to an intact group of both high and low criterion-setting eyewitnesses. Throughout this manuscript we use the labels low and high criterion-setting eyewitnesses to distinguish between those who are relatively willing and those who are relatively unwilling to choose. It is not clear that this pattern represents some stable individual difference phenomenon or merely something that afflicts participants on a particular identification task. We use these labels solely to reflect choosing differences on identification tasks and not with the intention of introducing a new individual difference variable to the literature.

The research on blank lineups suggests that choosing rates will decrease after the initial showup. Only those participants who reject the suspect presented in the initial showup – eyewitnesses with stringent criterion-setting – would be presented with a second showup. Accordingly, choosing should be lower in the second showup than in the first. It may be the case that choosing continues to decrease after the second showup, but we have no theoretical or empirical foundation upon which to make this prediction.

Unbiased Instructions

One way in which the cumulative risk associated with multiple showup procedures might be attenuated is through the use of unbiased instructions. Unbiased instructions are provided to the eyewitness immediately before an identification procedure and communicate to the eyewitness, at a minimum, that the perpetrator may not be present. Research examining the use of unbiased instructions in the context of lineup identification procedures indicates that innocent suspects are at less risk of

mistaken identification when eyewitnesses are instructed that the perpetrator may not be present (e.g., Clark, 2012; Steblay, 2013). While debate exists over the precise mechanism through which unbiased instructions impact eyewitness decision-making (e.g., Clark, 2005; Steblay, 2013), the pattern of results is consistent with a stringent shift in response criterion.

It is not completely clear, however, that instructing eyewitnesses that the perpetrator may not be present would be as effective in the context of a showup. In a lineup, an eyewitness is presented with multiple response options (suspect + fillers), so without explicitly telling eyewitnesses that the perpetrator may not be present, the eyewitness might assume that the perpetrator is present and that his or her job is simply to "pick him out". In contrast, simply asking if the suspect is the perpetrator in a showup actually implies that the suspect might not be the perpetrator. Accordingly, other components of unbiased lineup instructions (e.g., Technical Working Group for Eyewitness Evidence [TWGEYE], 1999) might be more important in showup procedures.

The TWGEYE (1999) recommended that, in addition to stating that the perpetrator may-or-may-not-be-present, administrating officers instruct eyewitnesses that clearing the innocent and identifying the guilty are equally important, and that police will continue to investigate the case regardless of whether the eyewitness makes an identification. Instructing eyewitnesses that clearing the innocent and identifying the guilty are equally important might affect how eyewitnesses weight the costs of failing to identify a perpetrator as compared to incorrectly identifying an innocent suspect.

Likewise, assuring the eyewitness that police will continue to investigate regardless of whether an identification is made might reduce any inherent demand to choose. We would expect that all three components of these unbiased instructions lead eyewitnesses to adopt more stringent response criteria. Though we do not contrast the different aspects of these instructions, we examine their potential, as a package, to reduce innocent suspect identifications over multiple showups (Study 4).

In summary, in this package of studies we: surveyed police officers on the use of multiple showups with the same eyewitness in cases involving a single perpetrator (Study 1); manipulated the number of innocent suspect showups that preceded a target showup (Study 2 and Study 3); and manipulated unbiased showup instructions (Study 4). Our aim was to examine the previously unexamined police practice of conducting multiple showups with the same eyewitness. We examined three research questions: (a) at what rate does the proportion of innocent suspect identifications increase relative to perpetrator identifications; (b) does response criterion become more stringent after the initial showup; and (c) how effective are unbiased showup instructions in multiple showup procedures?

Study 1: Do Police Use Multiple Showups?

The purpose of the first study was to determine if North American police officers were using multiple showups with the same eyewitness, how frequently they were doing so, and to gather a sense of the upper limit of showups they had conducted with the same eyewitness.

Methods

Participants

We collected survey data from a convenience sample of 284 police officers. Of those officers, 117 were from Canada, representing nine provinces and two territories. The remaining 167 officers were from the United States, representing 34 states.

Survey Instrument and Procedure

The information in this study was collected as part of a larger survey examining current police identification practices in Canada and the United States. Officers were instructed to answer based on the procedures they had used in the previous 12 months. All surveys were administered over the Internet.

Officers were recruited to participate in this study between January 2008 and January 2011. Individual officers were sent emails requesting their participation in the survey, while Chiefs of Police were emailed and sent post mail requests to have their officers participate in the survey. Contacts were provided with a link to the survey and were asked to distribute it to whomever they thought that it applied to.

Results

Of the 284 police officers that completed the survey, 97 (34.15%) reported that they had conducted a showup within the previous year. Those 97 officers were asked with regard to a case involving a single perpetrator, how often they had an eyewitness reject a suspect from a showup only to find a new suspect and present that suspect to the same eyewitness in a new showup. Given that only 13 (11.11%) of 117 officers from

Canada reported conducting a showup within the past year, we do not make inferential comparisons between Canadian and American police officers. Canadian responses were bimodal with 46.15% (n = 6) indicating that they had "never" conducted multiple showups with the same eyewitness and 46.15% (n = 6) indicating that they "sometimes" conducted multiple showups with the same eyewitness. The mode for American officers was "sometimes", with 44.05% of officers (n = 37) selecting this response. See Figure 1.

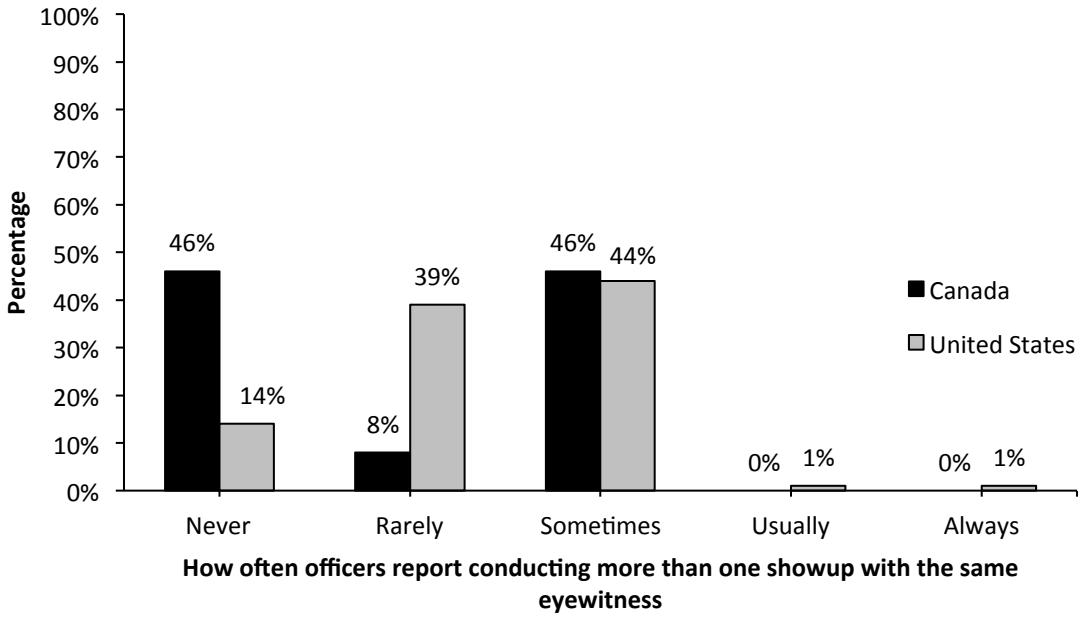

Figure 1. Regularity with which law enforcement personnel use multiple showup procedures

Note. These data are based on the 13 Canadian officers and 84 American officers who reported conducting a showup within the past year.

Next, officers were asked to report the highest number of showups they had conducted with any given eyewitness or a single perpetrator crime within the past year.

Of the 13 Canadian officers who reported conducting a showup within the past year, the average was 1.69 (*SD* = 0.61), while the average for the 84 American officers who reported conducting a showup within the past year was 4.43 (*SD* = 11.82). There were outliers among the American officers, however, with one officer claiming to have conducted 100 showups and another claiming to have conducted 50 showups with a single eyewitness seeking a single perpetrator. The median for American officers was 2. Notably, 19% of officers reported that the highest number of showups they conducted with the same eyewitness was four or greater. See Figure 2.

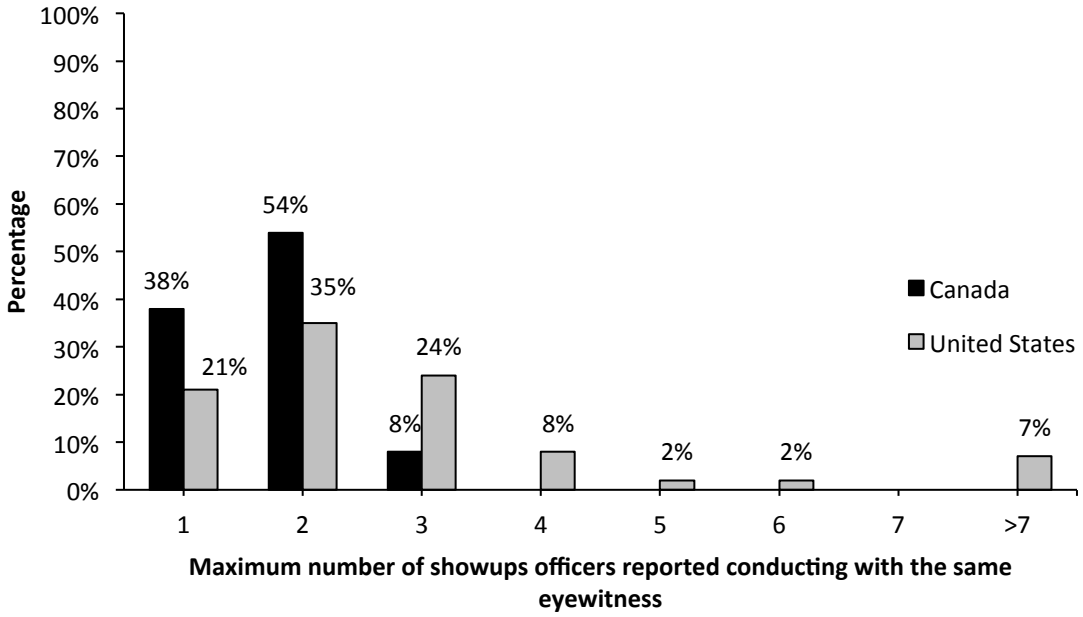

Figure 2. Highest number of showups law enforcement personnel reported using with the same eyewitness for a single-perpetrator case

Discussion

The percentage of officers who reported conducting showups within the past year was modest. This modest reporting might be the case because we sought to recruit officers who regularly conducted identification procedures, and asked gatekeepers to pass the survey along to these individuals. Our sampling may have been biased towards individuals who conduct lineup procedures, and, in Canada at least, detectives normally conduct lineup procedures, while constables (i.e., officers who are more active on patrol) are more likely to conduct showups. Regardless, the results of the survey suggest that it is not uncommon for North American officers to conduct multiple showups with the same eyewitness for single perpetrator crimes. Likewise, although the majority of officers report conducting no more than two or three showups with the same eyewitness, a substantial minority reported conducting four or more showups with the same eyewitness for a single perpetrator crime.

This survey highlights important concerns for eyewitness researchers. In the following three studies, we examined eyewitness identification performance over multiple showups.

Study 2: Establishing the Multiple Showup Effect

In this study, participants viewed a video of the target, followed by a showup. Each time participants rejected a showup they were returned to a filler task before being presented with a new showup. This process was repeated until participants made an identification or rejected all seven showups. We terminated the showup procedure after

an identification as the police would also stop after an identification. Similar to real-world situations, participants never knew how many showups they would (or could) see.

This experiment was exploratory, so we sought to maximize the effect size before examining more fine-grained effects (Study 3). Accordingly, participants were assigned to conditions in which zero, three, or six innocent suspect showups preceded the target. The main purpose of this experiment was to examine the rate at which innocent suspect identifications increased relative to perpetrator identifications, as measured by differences in a cumulative measure of innocence risk. We also examined differences in response bias across repeated showups.

Methods

Participants

We recruited undergraduates (N = 172) from the Psychology Subject Pool or through campus email in exchange for cash or course credit. We obtained demographic information for 156 (90.70%) of these participants – demographics were not collected for the remaining 16 participants due to a computer programming error. Participants on average were 18.86 (SD = 1.52) years of age and 80.13% of participants were female. The majority of participants were of European ancestry: European (74.36%), Asian (18.59%), African (0.64%), Middle Eastern (2.56%), First Nations (0.64%), and "other" (3.21%).

Materials

Targets

Although each participant saw only one target, we used three targets to attain some degree of stimulus generalization (Wells & Windschitl, 1999). For each target there were six innocent suspects who resembled the physical description of that target. All targets were male, European, in their early 20s, and varied in hair length and color (very short, light brown hair; medium-length, very dark hair; medium-length blonde hair).

Though we did not intentionally manipulate the extent to which innocent suspects resembled their respective targets, we presented 61 participants with each of the target photos next to each of their respective innocent suspects to examine whether similarity differences existed among the three groups. Participants were asked to use a slider to indicate how much innocent suspects resembled their respective targets from 0 (not at all) to 100 (exact match). We examined differences in the average similarities of innocent suspects to their respective targets. The extent to which innocent suspects resembled their respective targets, on average, did differ among the groups, Wilks $\lambda = .73$, $F(2, 59) = 11.03$, $p < .001$, partial $\eta^2 = .27$, with mean target-innocent suspect resemblance ratings ranging from 37.87 ($SE = 1.87$) to 46.17 ($SE = 1.89$). To examine whether the choosing pattern differed by target, we calculated a series of contingency tables examining the choosing pattern as a function of condition. For each of the contingency tables we removed one of the three targets and his respective innocents. The same choosing pattern

remained no matter which group was removed. Therefore, we collapsed all subsequent analyses across targets.

Each target video was a three-quarter angle of the target's face, in which the target made a brief statement related to criminal activity (e.g., "So, there's good stuff in that place?"). All videos were approximately 18 x 25 cm in size and were two seconds in duration. Each showup photo depicted only a face from the neck up (no clothing visible) looking straight into the camera. The background in each photo was constant (no context effects). Each photo was approximately 22 x 28 cm.

Filler task

The filler task involved the presentation of a complex beach scene taken from a "Where's Waldo?" book.6 Participants were given a sheet of paper with a list of questions regarding the image, such as "how many people in the picture are in red swimsuits?" The filler task was presented after the video of the target and after each time a participant rejected a showup. The length of the filler task between showup procedures varied from 20 to 210 seconds. These variations were designed to limit the ability of participants to determine precisely when (or if) another showup would be conducted. Although the durations varied among showups, they were fixed among participants (e.g., the length of time between the first and second showup was the same for all participants).

---

6 TM & © 2008 Entertainment Rights Distribution Limited. All rights reserved.

Procedure

Participants were brought into a room, seated at a computer, and instructed to watch a video. After watching the video, the letter of information appeared on screen and participants provided consent. After consent was given, participants completed the "Where's Waldo" filler task. The task was interrupted by the appearance of a face on screen, with the question "Do you believe this is the person you previously saw in the video?" We used a forced-choice decision task, so participants had the options of selecting "Yes" or "No" ("Don't Know" was not provided as a response option). After providing a response, each participant was asked to rate his or her confidence in the decision from 0 to 100%.

If participants selected "No", the Where's Waldo task resumed after the confidence rating, and this procedure continued until the participant either made an identification or rejected all seven showups. The procedure terminated once a participant identified a suspect – thus, the total number of showups a participant saw was dependent on the participants' identification decisions. Participants had no knowledge of how many showups they would (or could) see, nor did they know when a new face would be presented. When participants were presented with the first showup, they had no reason to expect that they might be presented with another showup. The order of suspect presentation within each condition was randomized with the exception that the target was fixed to a given position (the first, fourth, or seventh showup).

Design

Participants were randomly assigned in weighted blocks to one of three conditions created by placing the target showup in the first, fourth, or seventh position. Given that only a fraction of participants would survive to see the target in later showups, we assigned more participants to the conditions in which the target was presented in the fourth or seventh showup. Two in every 12 participants were assigned to the condition in which the target was presented in the first showup, five in every 12 participants were assigned to the condition in which the target was presented in the fourth showup, and the remaining participants were assigned to the condition in which the target was in the seventh showup.

Results

Hits

In Table 1 we present the frequency with which participants chose a suspect from a given showup for each of the three conditions. Using a binary logistic regression we examined whether hits (target identifications) differed as a function of the showup in which the target was presented. We used a repeated coding system for all hits and false alarm analyses in studies 2 and 3. With a repeated coding system, each level is compared to the subsequent level – performance when the target was in the first showup was compared with performance when the target was in the fourth showup, which in turn, was compared with performance when the target was in the seventh showup; however,

performance when the target was in the first showup was not compared with performance when the target was in the seventh showup.

Given that only 33.80% of participants survived to see the target when he was presented in the fourth showup and only 20.55% survived to see the target when he was presented in the seventh showup, we first examined the proportion of hits for only those participants who saw the target. Note that these analyses also represent differences in misses among the conditions, as misses are equal to the reciprocal of hits for those who saw the target. Target identifications occurred at a higher rate when the target was presented in the first showup (92.86%; $n = 28$) than when the target was presented in the fourth showup (41.67%; $n = 24$), $B = 2.90$, $SE = .84$, Wald's $\chi^2(1) = 11.86$, $p = .001$, $e^B = 18.20$ (95% CI [3.49, 94.90]). When the target was presented in the first showup, the odds that participants would identify the target (13:1) were 18.20 times greater than when the target was presented in the fourth showup (0.714:1). Participants were not significantly more likely to identify the target when he was in the fourth showup compared to when he was in the seventh showup (29.41%; $n = 17$), $B = 0.45$, $SE = .68$, Wald's $\chi^2(1) = .44$, $p = .51$, $e^B = 1.57$ (95% CI [0.41, 5.96]).

Table 1: *Frequency of Suspect Selections by Condition and Position in Study 2*

	Position							
	1	2	3	4	5	6	7	None
Target Position								
1 (*n* = 28)	**26**	0	0	0	0	0	0	2
	(93%)	(0%)	(0%)	(0%)	(0%)	(0%)	(0%)	
4 (*n* = 71)	31	8	8	**10**	2	1	1	10
	(44%)	(20%)	(25%)	**(42%)**	(14%)	(8%)	(9%)	
7 (*n* = 73)	28	6	8	8	5	3	**5**	10
	(38%)	(13%)	(21%)	(26%)	(22%)	(17%)	**(33%)**	
Totals:	85	14	16	18	7	4	6	22

Note. Conditional percentages are displayed in parentheses; i.e., % reflects percent of those reaching that showup who selected the person presented. Target identification frequencies and conditional percentages of those who identified the target are bolded and underlined.

Considering only those participants who saw the target does not capture the full extent of the problem. If the eyewitness identified an innocent suspect before seeing a showup containing the perpetrator, the eyewitness would never see the perpetrator, because police would not continue searching for suspects. Thus, we also examined the unconditional differences in target identifications. Participants were more likely to identify the target when he was presented in the first showup (92.86%; $n = 28$) compared to when he was presented in the fourth showup (14.08%; $n = 71$), $B = 4.37$, $SE = .81$, Wald's $\chi^2(1) = 29.21$, $p < .001$, $e^B = 79.30$ (95% CI [16.24, 387.34]). When the target was presented in the first showup, the odds that participants would identify him (13:1)

were 79.30 times greater than when the target was presented in the fourth showup (0.164:1). Participants were not significantly more likely to identify the target when he was in the fourth showup compared to when he was presented in the seventh showup (6.85%, n = 73), B = 0.80, SE = .58, Wald's $\chi^2(1) = 1.94$, $p = .16$, $e^B = 2.23$ (95% CI [0.72, 6.89]).

Cumulative False Alarms

Next we examined the impact of multiple showups on the cumulative false alarm rate (innocent suspect identifications). We estimated false alarms for a single showup by calculating the proportion of participants who chose the first (innocent) suspect with whom they were presented from the conditions in which the target was presented in the fourth or seventh showup. Likewise, to calculate the cumulative false alarm rate for the first four showups, we summed the proportions of participants who chose the first, second, third, or fourth (innocent) suspect with whom they were presented from the condition in which the target was presented in the seventh showup (the only condition in which participants could see four innocent suspects without having already seen the target). We could not calculate the false alarm rate for seven showups because we did not have a condition in which participants could see an innocent suspect in the seventh showup. Therefore, we calculated the false alarm rate for the first six showups from the condition in which the target was presented in the seventh showup (a slightly conservative estimate).

Not surprisingly, the proportion of false alarms for the first showup was lower (40.97%) than the cumulative proportion of false alarms for the first four showups (68.49%), B = -1.14, SE = .30, Wald's $\chi^2(1) = 14.14$, $p < .001$, $e^B = 0.32$ (95% CI [0.18, 0.58]). The odds that participants would commit a false alarm (0.69:1) when presented with only one showup were 68% less likely (0.32 − 1= -0.68 × 100 = - 68%) than when participants were presented with up to four showups (2.17:1). Participants did not commit significantly fewer false alarms in the first four showups than they did in the first six showups (79.45%), B = -0.58, SE = .38, Wald's $\chi^2(1) = 2.25$, $p = .13$, $e^B = 0.56$ (95% CI [0.27, 1.19]).

Confidence

We examined the relationship between accuracy and confidence for both identifications from the first showup and identifications from the fourth showup. Given that we did not have a condition in which participants could view an innocent suspect in the seventh showup without having already rejected the target, we did not examine the relationship between confidence and accuracy for the seventh showup. On the first showup, participants who accurately identified the target were more confident (M = 76.96, SD = 13.87) than those who identified an innocent suspect (M = 67.71, SD = 15.84), $t(83) = 2.57$, $p = .01$, $r = .27$; however, on the fourth showup, those who accurately identified a target were no more confident (M = 63.40, SD = 20.21) than those who identified an innocent suspect (M = 56.14, SD = 18.28), $t(15) = 0.76$, $p = .46$, r = .19.

Given that only two participants who saw the target in the first showup rejected him and only four participants who saw the target in the fourth showup rejected him, we do not examine the relationship between confidence and accuracy for rejections in this experiment.

Innocence Risk

Next, we calculated innocence risk (Clark & Godfrey, 2009) for one, four, and seven showups. To do so, we divided the proportion of false alarms by the sum of false alarms and hits. We adjusted innocence risk to reflect the cumulative false alarm rates associated with conducting multiple showups, and the fact that not all eyewitnesses survived to see the target. We used the false alarm rates and the unconditional hit rates calculated above to reflect the cumulative risk. As mentioned in our discussion of false alarms, we could only calculate the false alarm rate for a maximum of six showups, as we did not have a condition in which participants could see seven showups without having already seen the target. Accordingly, we underestimated the innocence risk associated with seven showups. The innocence risk values were .31, .83, and .92 when the target was presented in the first, fourth, and seventh showups, respectively. When considering only a single showup, there was a 31% chance that the suspect was innocent, given that he was identified. When considering the first four showups (with the target placed in the fourth showup), there was an 83% chance that the suspect was innocent, given that he was identified. When considering all seven showups (with the target placed in the

seventh showup), there was a 92% chance that the suspect was innocent given that he was identified.

Next, we used the modified jackknife procedure (Mosteller & Tukey, 1968; Horry, Palmer, & Brewer, 2012) to estimate the standard error of the innocence risk values. We then used the estimated standard errors to calculate inferential 95% confidence intervals (CIs) around each statistic (Tryon, 2001). Innocence risk was significantly lower in the first showup (.31, jackknife *SE* = .02, 95% CI = .28, .34) than in the first four (innocence risk = .83, jackknife *SE* = .04, 95% CI = .77, .89) or seven (innocence risk = .92, jackknife *SE* = .03, 95% CI = .87, .97) showups; however, there was no significant difference in innocence risk between four and seven showups.

Decision Criteria

Finally, we examined whether response bias was different between the first and fourth showups with which participants were presented. We could not examine response bias for the seventh showup because we did not have an adequate control. To compare decision criteria between the first and fourth showups, we used the same jackknife procedure (Mosteller & Tukey, 1968) to estimate standard errors and inferential 95% confidence intervals (Tryon, 2001) as we used to compare innocence risk. We used *c* as our measure of response bias. To calculate *c*, we converted the proportions of hits and false alarms to *z* scores, found the average, and multiplied the product by -1 (Stanislaw & Todorov, 1999). Therefore, negative values indicate a bias towards choosing, positive values indicate a bias towards not choosing, and a value of zero indicates no bias. As

predicted, response bias was more lax in the first showup with which participants were presented, c = -0.62, jackknife SE = .24, 95% CI = -1.04, -0.20, as compared to the fourth showup with which participants were presented, c = 0.45, jackknife SE = .18, 95% CI = 0.12, 0.80.

Discussion

As hypothesized, innocence risk was substantially higher through four showups than through the first showup. The effect of presenting innocent suspects in showups before presenting the target was devastating to identification performance. Surprisingly, there was no significant difference between conditions in which the target was presented in the fourth or seventh showup. This is likely a product of the fact that innocence risk was already very high, so there was little room for it to increase further. Furthermore, the proportion of hits decreased from the first to the fourth showup. Even when considering only those participants who survived to see the target, participants were more likely to identify the target in the first showup than they were in the fourth or seventh showups.

This study had several limitations. First, hits were near ceiling when the target was presented in the first showup. This was likely a result of the short memory retention interval between encoding and the first recognition trial – approximately three and a half minutes. The high rate of hits when the target was in the first showup probably led to an exaggerated effect size when comparing this condition to conditions in which the target was presented in either the fourth or seventh showup. The ceiling effect for hits when the

target was presented in the first showup increased the proportion of identified suspects that were guilty and may have led to an underestimation of innocence risk.

Second, the target was presented in the first, fourth, or seventh showup as opposed to the first, second, or third showup. This first experiment was exploratory, so we presented the target in the fourth or seventh showup because, although we wanted to maximize the effect, we did not expect the magnitude of effect to be this large. Moreover, as our first study indicates, 19% of officers reported conducting four or more showups with the same eyewitness. This study provides compelling evidence that conducting this many showups with the same eyewitness is both dangerous to innocent suspects, and limited in terms of the inferences that one can make if an eyewitness makes an identification from such a procedure – because innocence risk is high.

Third, although we found support for our hypothesis that decision criteria would be most lax in the initial showup, these data do not allow us to determine if the criterion shift was gradual or abrupt, as we did not have a condition in which participants could identify the target in the second showup. The blank lineup literature suggests that response bias should be more stringent in the second showup relative to the first, but provides no information as to whether response bias should continue to become more stringent after the second showup.

Fourth, we could not properly estimate innocence risk for the seventh showup as we did not have a condition in which participants could see a seventh innocent suspect. Accordingly, we were forced to estimate innocence risk for seven showups using the

proportion of target identifications from the condition in which the target was presented in the seventh showup and using the proportion of innocent suspect identifications from six innocent suspect showups.

Despite the numerous limitations of this study, we believe that it provides compelling evidence that conducting four or more showups with the same eyewitness for a single perpetrator crime increases innocence risk to unreasonably high levels. Given that conducting four showups with the same eyewitness produced an innocence risk value of .83, it is difficult to argue that this procedure produces evidence with sufficient probative value to be useful in court. Innocence risk values in excess of .50 indicate that a suspect is more likely to be innocent than guilty, given identification. Although effect sizes will vary from one study to another, the magnitude of this effect makes it difficult to envision a procedure that would make it reasonable to conduct this many showups with the same eyewitness.

Study 3: Fine-Grained Effects of Multiple Showups

The goal of the third study was to address the limitations of the second study and to examine more fine-grained effects. Accordingly, we (1) increased the memory retention interval between encoding and the first showup, and (2) presented zero, one, two, three, or four innocent suspects before the target. The memory retention interval was increased in an attempt to eliminate the ceiling effect when the target was presented in the first showup, and to more closely reflect the retention intervals in actual cases. Presenting zero, one, two, three, or four innocent suspects before the target enabled us to

examine more fine-grained effects and allowed us to determine where the stringent shift in response bias was occurring (e.g., does the effect emerge between the first and second showup or does the response bias gradually become more conservative from the first to the fourth showup?). This design also allowed us to properly estimate innocence risk for the first four showups. The condition in which the target was presented in the fifth showup was included solely to enable us to accurately estimate innocence risk for the first four showups, by giving us a condition in which participants could see four innocent suspect showups before seeing the target. The procedure was largely the same as that in Study 2, with the exception of the differences mentioned in the Methods section. Based on the results of Study 2, we hypothesized that innocence risk would increase as the number of innocent suspect showups before the target showup increased, and, based on the blank lineup literature, we expected a more stringent response bias in the second showup relative to the first.

Methods

Participants

Undergraduates (N = 222) recruited from the Psychology Subject Pool or through campus email participated in this study in exchange for cash or course credit. Participants on average were 18.46 (SD = 4.12) years of age and 89.64% were female. The majority of participants were of European ancestry: European (71.17%), Asian (18.02%), African (1.35%), Middle Eastern (2.70%), First Nations (0.90%), "other" (4.50%). The remaining 1.35% did not disclose their ancestry.

Materials

Targets

We used the same materials as in Study 2, with the exception that we dropped two of the six innocent suspects for each of the respective targets. Selection of who was dropped was arbitrary and not based on similarity to suspect. As in Study 2, the choosing pattern in this study did not vary as a function of the target. Accordingly, we collapsed all analyses across targets.

Filler task

We also used a different filler task in Study 3. We had participants complete an anagrams task. Participants were required to rearrange the letters of nonwords to form words (e.g., "TABO" can be rearranged into the word "BOAT"). Participants were presented with the anagrams task before each showup. Participants were instructed to complete as many anagrams as they could and that if they could not solve one, they could press enter to skip ahead to the next anagram. We varied the length of the anagram task between showups from 30 to 300 seconds to keep participants from predicting when they might see another showup. Although the length of the anagrams task varied, anagram task duration was fixed among participants (e.g., the length of the anagrams task between showups one and two was the same for all participants).

Procedure

The procedure was the same as in Study 2 with a few notable exceptions. First, after participants viewed the target video, gave consent, and completed a demographic

questionnaire, they completed an unrelated study that took approximately 20 minutes. Second, as described above, we used a different filler task. Third, participants were assigned to a condition in which the target was in the first, second, third, fourth, or fifth showup; however, as in Study 2, the procedure stopped after participants made an identification or rejected all five showups – participants only saw the target if they rejected all innocent suspects presented before him.

Design

Participants were randomly assigned to conditions in which the target was presented in the first, second, third, fourth, or fifth showup.

Results

Hits

In Table 2 we provide the frequencies with which participants chose a suspect from a given showup for each of the five conditions. We first examined the conditional differences in hits. Participants were more likely to accurately identify the target when he was presented in the first showup (71.74%) than when he was presented in the second showup (42.11%), $B = 1.25$, $SE = .57$, Wald's $\chi^2(1) = 4.84$, $p = .03$, $e^B = 3.49$ (95% CI [1.15, 10.64]). The odds of a target identification were 3.49 times greater when the target was presented in the first showup (2.54:1) than when the target was presented in the second showup (0.73:1). No other significant differences emerged, Bs = - 0.69 to -0.12, $SEs \geq .65$, Wald's $\chi^2(1) \leq 0.61$, $p \geq .43$, e^Bs = 0.50 to 0.89.

Next, we examined differences in the unconditional proportion of hits. Participants were more likely to accurately identify the target when he was presented in the first showup (71.74%) than when he was presented in the second showup (20.00%), $B = 2.32$, $SE = .51$, Wald's $\chi^2(1) = 20.39$, $p < .001$, $e^B = 10.15$ (95% CI [3.71, 27.77]). The odds of a target identification were 10.15 times greater when the target was presented in the first showup (2.54:1) than when the target was presented in the second showup (0.25:1). No other significant differences emerged, Bs = 0.00 to 0.33, SEs $\geq .54$, Wald's $\chi^2(1) \leq 0.36$, ps $\geq .55$, e^Bs = 1.00 to 1.39.

Table 2: *Frequency of Suspect Identifications by Condition and Position in Study 3*

	Position					
	1	2	3	4	5	None
Target Position						
1 (*n* = 46)	**33**	2	4	1	0	6
	(72%)	(15%)	(36%)	(14%)	(0.00%)	
2 (*n* = 40)	21	**8**	4	1	1	5
	(53%)	**(42%)**	(36%)	(14%)	(17%)	
3 (*n* = 45)	18	7	**9**	1	3	7
	(40%)	(26%)	**(45%)**	(9%)	(30%)	
4 (*n* = 46)	18	7	7	**7**	1	6
	(39%)	(25%)	(33%)	**(50%)**	(14%)	
5 (*n* = 45)	26	6	2	2	**6**	3
	(58%)	(32%)	(15%)	(18%)	**(67%)**	
Totals:	116	30	26	12	11	27

Note. Conditional percentages are displayed in parentheses; i.e., % reflects percent of those reaching that showup who selected the person presented. Target identification frequencies and conditional percentages of those who identified the target are bolded and underlined.

Cumulative False Alarms

Next we examined the impact of multiple showups on the cumulative false alarm rate. We used the same process described in Study 2 to estimate the cumulative false alarm rate for one, two, three, and four showups. To calculate the proportion of false alarms for a given number of showups we used data from all conditions in which the

target was presented in a later showup. For example, to compute the false alarm rate for two showups, we calculated the proportion of participants who made false alarms in either the first or second showup from the conditions in which the target was presented in the third, fourth, or fifth showup.

Not surprisingly, the false alarm rate was lower in the first showup (47.16%) than in the first two showups (60.29%), B = -0.53, SE = .23, Wald's $\chi^2(1) = 5.28$, p = .02, e^B = 0.59 (95% CI [0.37, 0.93]) and marginally lower in the first two showups than in the first three showups (73%), B = -0.55, SE = .29, Wald's $\chi^2(1) = 3.56$, p = .06, e^B = 0.58 (95% CI [0.32, 1.02]); however, the false alarm rate was not significantly lower through the first three showups than it was through the first four showups (80%), B = -0.42, SE = .44, Wald's $\chi^2(1) = 0.89$, p = .35, e^B = 0.66 (95% CI [0.28, 1.57]). When participants were presented with only one innocent suspect showup (0.89:1) they were 41% (0.59 − 1= -0.41 × 100 = -41%) less likely to commit a false alarm than when participants were presented with up to two innocent suspects (1.52:1); likewise, when participants were presented with up to two innocent suspects, they were 42% (0.58 − 1 = -0.42 × 100 = -42%) less likely to commit a false alarm than when participants were presented with up to three innocent suspects.

Confidence

Means, standard deviations, sample sizes, *t*s, and effect sizes are displayed in Table 3 for both selections and rejections. We found no relationship between confidence and accuracy for selections on any of the showups. Correct rejections of innocent

suspects on the first showup were made with greater confidence (M = 56.96, SD = 23.02) than incorrect rejections of the target (M = 43.23, SD = 20.61), $t(104) = 2.04$, $p = .04$, $r = .20$. Likewise, correct rejections of innocent suspects on the second showup were made with greater confidence (M = 68.59, SD = 22.58) than incorrect rejections of the target (M = 48.55, SD = 31.09), $t(63) = 2.51$, $p = .02$, $r = .30$; however, no other significant differences emerged.

Table 3: *Relationship Between Confidence and Accuracy by Showup for Study 3*

	Confidence Selections								
	Accurate			Inaccurate					
	M	*SD*	*n*	*M*	*SD*	*n*	*t*	*p*	*r*
Showup 1	65.00	19.10	33	64.31	17.01	83	0.19	.85	.02
Showup 2	49.38	21.91	8	59.79	16.80	19	-1.35	.19	-.26
Showup 3	63.67	23.39	9	69.56	15.83	9	-0.63	.54	-.15
Showup 4	66.29	16.70	7	70.00	28.28	2	-0.25	.81	-.09
	Confidence Rejections								
	Accurate			Inaccurate					
	M	*SD*	*n*	*M*	*SD*	*n*	*t*	*p*	*r*
Showup 1	56.96	23.02	93	43.23	20.61	13	2.04	.044	.20
Showup 2	68.59	22.58	54	48.55	31.09	11	2.51	.015	.30
Showup 3	72.12	34.12	25	62.91	29.53	11	0.78	.44	.13
Showup 4	56.11	27.24	9	70.71	26.84	7	1.07	.30	-.28

Innocence Risk

As in Study 2, we compared differences in innocence risk as a function of the number of innocent suspect showups seen before the target showup. As predicted, innocence risk values were significantly lower in the first showup (.40, jackknife *SE* = .03, 95% CI = .35, .44) than in the first two showups (.75, jackknife *SE* = .06, 95% CI = .66, .84). No other significant differences emerged. Innocence risk values, standard errors, and confidence intervals for all conditions are provided in Table 4.

Table 4: *Innocence Risk as a Function of the Number of Innocent Suspect Showups Shown Before the Target Showup in Study 3*

		Jackknife Statistics		
Target Showup	Innocence Risk	*M*	*SE*	95% CI
Showup 1	.40	.40	.03	.35 - .44
Showup 2	.75	.75	.06	.66 - .84
Showup 3	.78	.78	.05	.71 - .86
Showup 4	.84	.84	.05	.77 - .91

Decision Criteria

Finally, we compared response bias among the showups. As predicted, response bias was more lax in the initial showup with which participants were presented (*c* = -0.25, jackknife *SE* = .11, 95% CI = -0.41, -0.09) as compared to the second showup with which participants were presented (*c* = 0.41, jackknife *SE* = .17, 95% CI = 0.16, 0.65).

Response bias did not become any more stringent with additional showups. Response bias, jackknife standard errors, and 95% CIs for each showup are summarized in Table 5.

Table 5: *Decision Criteria by Showup for Study 3*

		Jackknife Statistics		
	c	*M*	*SE*	95% CI
Showup 1	-0.25	-0.25	0.11	-0.41 - -0.09
Showup 2	0.41	0.41	0.17	0.16 - 0.65
Showup 3	0.42	0.42	0.19	0.15 - 0.69
Showup 4	0.45	0.45	0.33	-0.03 - 0.94

Note. Negative values indicate lenient decision criteria and positive values indicate stringent decision criteria. In addition, the wide confidence interval in the fourth showup is due to the small sample of participants that survived to the fourth showup.

Discussion

Overall, we found support for both of our hypotheses. Consistent with Study 2, innocence risk was greater when participants were presented with more than one showup. We also found support for our hypothesis that response bias would be more stringent on the second showup than it was on the first. Response bias did not become any more stringent after the second showup.

We failed to replicate the relationship between confidence and accuracy for selections that we found in Study 2. Yet, we did find a confidence-accuracy relationship for rejections. It is difficult for us to make sensible conclusions about the presence and absence of the significant correlations between confidence and accuracy, as the package

of studies in the present manuscript were not designed to systematically address this issue.

Study 3 overcame several of the limitations in Study 2. First, in Study 2, hits were approaching ceiling (93%) when the target was presented in the first showup. In Study 3, increasing the memory retention interval brought the hit rate in the first showup down to 72%. This allowed us to more accurately estimate innocence risk in the first showup, which enabled us to better gauge the effect of conducting multiple showups on innocence risk. Second, in this study we were able to examine more fine-grained effects of conducting multiple showups with the same eyewitness. Indeed, we found that the use of even two showups with the same eyewitness led to dangerously high levels of innocence risk. Finally, response bias was more stringent on the second showup compared to the first, suggesting that low criterion-setting eyewitnesses were filtered out on the initial showup. Response bias became no more stringent on subsequent showups.

We believe that Studies 2 and 3 provide compelling evidence that the use of multiple showups is a dangerous procedure in regards to its evidentiary value, raising the question of whether there is a way to decrease the innocence risk associated with the use of multiple showups. We explored this prospect in Study 4.

Study 4: Effects of Unbiased Showup Instructions

The goal of our fourth study was to examine the potential for unbiased instructions to drive down the false alarm rate and, thus, increase the proportion of participants that "survive" to see a suspect in a given showup. Past research indicates

that unbiased lineup instructions reduce innocent suspect identifications (Clark, 2012; Steblay, 2013). In this study, we sought to examine how many innocent suspect showups participants could endure before committing a false alarm when they received unbiased showup instructions before each showup. We compared this group to two control groups; one received no instructions and the other received instructions before the first showup only. If police conducted a series of showups with the same eyewitness over a relatively short period of time, they may find it redundant to repeat unbiased instructions before each showup. Therefore, we included the unbiased instructions before the first showup only as a control group to examine how providing unbiased instructions before the first showup only would impact eyewitness decision-making.

Participants were presented with up to four innocent suspect showups, and if they rejected all innocent suspect showups, a fifth showup that included the target.

Methods

Participants

We recruited undergraduates (N = 184) from the Psychology Subject Pool and through campus email in exchange for cash or course credit. Participants, on average, were 18.64 (SD = 2.89) years of age and 83.15% of participants were female. The majority of participants were of European ancestry: European (77.72%), Asian (13.04%), African (2.72%), Middle Eastern (1.09%), "other" (5.43%).

Materials

We used the same materials as in Study 3. As in studies two and three, the choosing pattern did not vary as a function of the targets. Accordingly, we collapsed all analyses across the three targets. The filler task was the same as in Study 3.

Procedure

Participants were brought into a room, seated at a computer, and instructed to watch a video. Immediately after watching the video, participants were told that if this were a real crime, police would begin searching for a suspect and might present them with a suspect at a later time. Afterwards, the letter of information appeared on screen. Consent to participate was obtained at this time. Participants then provided their age, ethnicity, and biological sex before completing an unrelated study that lasted about 30 minutes. Next, participants completed the same anagrams task that we used in Study 3. After the allotted time for the first series of anagrams had elapsed, participants in two of the three conditions (i.e., those receiving instructions at least once) were presented with the unbiased showup instructions. Specifically, the instructions stated:

> In a moment, you will view a photo of a man that may or may not be the man you saw in the video. If you believe that the man in the photo is the man you saw in the video, you should identify him as such (click "YES"). If you do not believe that the man in the photo is the man you saw in the video, you should not identify him (click "NO"). Remember, it is just as important to clear innocent persons from suspicion as it is to identify guilty parties. Regardless of whether you make an identification, police will continue to investigate the incident.

These instructions were designed to mimic those recommended by the TWGEYE (1999). Next, all participants – including those who did not receive unbiased showup instructions – were presented with the first showup. The procedure was the same as in Study 3, with the exceptions of our instruction manipulation, and that the target was always in the fifth showup. If participants correctly rejected all four innocent suspect showups, they were presented with a fifth showup that contained the target.

Design

Participants were randomly assigned to one of three conditions in which they received unbiased instructions before each showup, unbiased instructions before the first showup only, or no unbiased instructions.

Results

Survival

First, we calculated the proportion of participants that survived to see the target as a function of the instructions. The proportion of eyewitnesses making a selection on a given showup is summarized for each condition in Table 6. Participants were more likely to survive to see the target when they received instructions before the first showup only (54.10%), $B = 0.86$, $SE = .37$, Wald's $\chi^2(1) = 5.35$, $p = .02$, $e^B = 2.36$ (95% CI [1.14, 4.88]), or before each showup (51.67%), $B = 0.76$, $SE = .37$, Wald's $\chi^2(1) = 4.18$, $p = .04$, $e^B = 2.14$ (95% CI [1.03, 4.43]), compared to when they did not receive any instructions (33.33%); however, there was no statistically significant difference between those who received instructions before the first showup only and those who received instructions

before each showup, $B = 0.10$, $SE = .36$, Wald's $\chi^2(1) = 0.07$, $p = .79$, $e^B = 1.10$ (95% CI [0.54, 2.25]). Participants who were provided with instructions before the first showup only (1.18:1) or before each showup (1.07:1) were 2.36 and 2.14 times as likely to survive to see the target, respectively, than were participants who did not receive instructions (0.50:1).

Next, we examined mean differences in the number of correctly rejected showups as a function of instructions. Instructions significantly influenced the mean number of showups rejected by participants, $F(2, 181) = 4.34$, $p = .01$, partial $\eta^2 = .05$. Participants provided with instructions before the first showup only, on average, rejected more showups ($M = 2.64$, $SD = 1.62$) than participants not provided with any instructions ($M = 1.89$, $SD = 1.71$), $t(122) = 2.50$, $p = .01$, $d = 0.45$. Likewise, participants who received instructions before each showup rejected more showups ($M = 2.65$, $SD = 1.61$) than did participants not provided with any instructions, $t(121) = 2.53$, $p = .01$, $d = 0.46$; however, there was no statistically significant difference between those participants who received instructions before the first showup only and those who received instructions before each showup, $t(119) = -0.04$, $p = .97$, $d = 0.01$.

Table 6: *Frequency of Suspect Selections by Condition and Position for Study 4*

	1	2	3	4	5	None
Instructions						
None	22	8	9	3	**5**	16
(*n* = 63)	(35%)	(20%)	(27%)	(13%)	**(24%)**	
Before first showup	11	6	10	1	**15**	18
(*n* = 61)	(18%)	(12%)	(23%)	(3%)	**(45%)**	
Before all showups	11	6	7	5	**13**	18
3 (*n* = 60)	(18%)	(12%)	(16%)	(14%)	**(42%)**	
Totals:	44	20	26	9	33	52

Note. Conditional percentages are displayed in parentheses. Target identification frequencies and conditional percentages of target identifications are bolded and underlined.

We also examined differences in correct rejections on the first showup. If unbiased instructions increase the proportion of eyewitnesses who correctly reject the initial showup, conducting two showups with the same eyewitness for a single perpetrator crime might be a viable option in some instances. Given that the unbiased instruction conditions were equivalent on the first showup, we collapsed these conditions for the purpose of this analysis. Participants who received unbiased instructions were

significantly more likely to reject the first innocent suspect with whom they were presented (82%) than were participants who did not receive unbiased instructions (65%), $B = 0.88$, $SE = .35$, Wald's $\chi^2(1) = 6.20$, $p = .01$, $e^B = 2.42$ (95% CI [1.21, 4.83]). Participants who were provided with unbiased instructions (4.50:1) were 2.42 times as likely to reject the first suspect than were participants who were not provided with unbiased instructions (1.86:1).

Hits

For participants who saw the target, the proportion of hits did not differ as a function of the instructions manipulation, Bs = 0.84 to 0.98, SEs $\geq .62$, Wald's $\chi^2(1) \leq 2.50$, ps $\geq .11$, e^Bs = 2.31 to 2.67. When examining the unconditional proportion of hits, participants were more likely to identify the target when they received instructions before the first showup only (24.59%), $B = 1.33$, $SE = .55$, Wald's $\chi^2(1) = 5.79$, $p = .02$, $e^B = 3.78$ (95% CI [1.28, 11.18]), or before each showup (21.67%), $B = 1.17$, $SE = .56$, Wald's $\chi^2(1) = 4.31$, $p = .04$, $e^B = 3.21$ (95% CI [1.07, 9.65]), than when participants did not receive any instructions (7.94%); however, there was no statistically significant difference between participants who received instructions before the first showup only and those who received instructions before each showup, $B = 0.17$, $SE = .43$, Wald's $\chi^2(1) = 0.15$, $p = .70$, $e^B = 1.18$ (95% CI [0.51, 2.75]). The odds of a target identification were 3.78 times greater for those participants who were presented with instructions before the first showup only (0.33:1) and 3.21 times greater for those participants who

were presented with instructions before each showup (0.28:1) compared to those participants who were not provided with instructions (0.09:1).

Discussion

As predicted, participants receiving unbiased showup instructions rejected more innocent suspects and were more likely to see the target than those participants who did not receive unbiased showup instructions. Moreover, greater proportions of those who received instructions identified the target. Although instructions improved both the proportion of participants surviving to see the target and the proportion of target identifications in a relative sense, performance in absolute terms was still poor. Of those participants receiving instructions before the first showup only or before each showup, 24.59% and 21.67% identified the target, respectively. Even when provided with instructions, only half survived to see the target, and less than half of those who saw the target identified him. This provides further evidence that, even under relatively ideal presentation conditions, using multiple showups with the same eyewitness prompts unreliable decisions with potentially dangerous consequences.

We also examined differences in the proportion of participants who correctly rejected the first showup as a function of the unbiased instructions. Instructions increased the proportion of participants correctly rejecting the first showup from 65% to 82%. Thus, while conducting more than a single showup with the same eyewitness is still dangerous, it may be tenable to conduct two showups with the same eyewitness so long as the eyewitness is provided with unbiased instructions. Unfortunately, we did not

include a condition in which participants could identify the target in the second showup. Accordingly, we cannot calculate innocence risk in order to better assess the viability of conducting two showups with the same eyewitness.

General Discussion

Results from our first study showed that police do sometimes conduct multiple showups with the same eyewitness, and that a substantial minority of police officers will use four or more showups with the same eyewitness. Further, results showed that response bias was more stringent after the initial showup (Study 2) but became no more stringent after the second showup (Study 3). Despite response bias becoming more stringent, false alarms increased with additional showups, as they are cumulative. Innocence risk also increased, reflecting that innocent suspect identifications increased while perpetrator identifications decreased. Unbiased showup instructions attenuated the deleterious effects of using multiple showups with the same eyewitness, but the outcome was still less than optimal. Innocence risk was relatively high for even a single showup (Study 2 = .31, Study 3 = .40), but when conducting multiple showups, innocence risk was in excess of 50% (Studies 2 and 3) meaning that, a suspect was more likely to be innocent than guilty, given identification. These studies provide compelling evidence that using multiple showups with the same eyewitness will produce an unacceptable number of innocent suspect identifications.

Theoretical Implications

Clearly, the more stringent response bias in the second showup compared to the first is due, in part, to the attrition of low criterion-setting eyewitnesses; however, there are at least two other mechanisms that also might be operative. First, presenting eyewitnesses who reject an initial showup with a second showup might affect their perceptions of scarcity. On the initial showup, eyewitnesses have no reason to expect that they might have additional opportunities to identify the perpetrator – accordingly, this showup might be viewed as the only opportunity to identify the perpetrator. In an attempt to seize the opportunity, an eyewitness might adopt a lax decision strategy; however, when police present the eyewitness who rejects the initial showup with a second showup this conveys to the eyewitness that opportunities to identify the perpetrator are less scarce than the eyewitness once perceived. Given this newly acquired information, the eyewitness can now afford to adopt a more stringent decision criterion.

Some credence for this potential mechanism comes from research examining sequential lineups, which indicates that eyewitnesses gradually adopt more lenient decision criteria as the lineup progresses (Horry et al., 2012). One interpretation of this effect is that, as the end of a lineup approaches, an eyewitness realizes that opportunities to identify the perpetrator are becoming increasingly scarce. In response, the eyewitness adopts a more lax decision criterion to maximize his or her likelihood of identifying the perpetrator. In the case of multiple showups, the criterion moves in the opposite direction

because participants now perceive opportunities to identify the perpetrator as less, as opposed to more, scarce.

A second mechanism that might be operating in the multiple showup procedure is base-rate neglect. During the initial showup, an eyewitness might neglect to consider the probability that the suspect is the target. Indeed, when making judgments, people often rely on individuating information to the exclusion of base-rate information (e.g., Kahneman & Tversky, 1973). Yet, individuals will use base-rate information to the extent that it communicates causal implications for the outcome (Ajzen, 1977). Presenting an innocent suspect in the first showup provides no information about the likelihood of perpetrator-presence on a second showup, but it does convey important base-rate information about how often police suspect the correct person – *not always*. Having just been presented with the wrong person, an eyewitness might be less optimistic that a suspect presented in a second showup is the perpetrator, and therefore the eyewitness might adopt a more stringent decision criterion.

Both the perceptions of scarcity and base-rate accounts predict a more stringent response bias on the second, compared to the first, showup. We cannot tease apart these two mechanisms with the previously presented studies, and there is no reason to assume that the two mechanisms are mutually exclusive. Future research might examine the role of each mechanism through administrating pre- and post-test questionnaires tapping perceptions of base-rates and opportunities. Due to the demand that might be associated with a pre-test questionnaire, it would not be wise to include this in the same

experimental session as the multiple showups procedure or post-test questionnaire, but it could be administered months prior, perhaps during participant screening. The difference vectors could then be included in a regression model as predictors of choosing from a multiple showups procedure. Further, adding a no pre-test control might provide a fail-safe in the instance that the pre-test affects participant responding on the initial showup.

Applied Implications and Recommendations

Practically, the different theoretical mechanisms do not alter the conclusion that the use of multiple showups with the same eyewitness elevates innocence risk values to dangerously high levels (>.50). Though unbiased showup instructions attenuated this effect, large proportions of participants still chose innocent suspects from the first or second showup.

Over the long run, the use of multiple showups with a single eyewitness will produce far too many innocent suspect identifications. Yet, if police find another suspect after an eyewitness has already rejected one or more suspects, it seems impractical to recommend that police cannot use another showup with the same eyewitness. Such a policy would surely stymy police investigations and allow perpetrators to walk free. Given the practical constraints of police investigations, multiple showups might be used as an investigative procedure, but, because of their low evidentiary value, an identification under such circumstances should not be treated as evidence of guilt. Instead, if an eyewitness makes an identification after being presented with multiple showup procedures, the police should use this as an indicator that they ought to

investigate this suspect further. In other words, we recommend that police use an identification from a multiple showup procedure as a heuristic that they should further investigate the person of interest, but the identification should not be used as evidence *per se*.

If police were to use a multiple showup procedure for investigative purposes, they would need to obtain evidence of guilt by means other than eyewitness identification in order to prosecute that suspect. It is not sufficient to use a properly constructed lineup procedure after an identification from a showup procedure because the reliability associated with a properly constructed lineup does not hold after the eyewitness has already identified the suspect from a showup (Haw, Dickinson, Meissner, 2007; Valentine, Davis, Memon, & Roberts, 2012). Attempts to obtain eyewitness evidence from more reliable eyewitness procedures are tainted by the suggestiveness of previous procedures (Dysart, Lindsay, Hammond & Dupuis, 2001).

When multiple eyewitnesses are available, however, police may use multiple showups with one eyewitness, and pending an identification, properly constructed lineups with other (untainted) eyewitnesses (e.g., TWGEYE, 1999). This practice is clearly beneficial in that it offers both the potential to make a swift arrest of the potential perpetrator and to obtain strong evidentiary value from a properly constructed lineup procedure.

General Limitations

There are two limitations of this research program that are particularly important to consider. First, our showups involved presenting eyewitnesses with photographs of suspects as opposed to the actual suspects themselves. Though police will use photographic showup procedures, it is considerably more common that police will use live showup procedures (Behrman & Davey, 2001). It would not have been practical for us to use live showups given the amount of resources this would have required. Moreover, the photographs we used included only the suspect's face and neck; the suspects' bodies were cropped and the background in all showups was constant – and different from that of the target video – to limit any contextual cues. Thus, the photographs we did use were low in ecological validity. Future research might examine cumulative innocence risk for two or more live showups.

Our results are also limited by the short delays between showups. With modern technology, police could present showups as quickly as participants in our study were presented with showups; however, it is reasonable to assume that the delays between showups would generally be longer than they were in our studies. The short delays between showups in our studies may have exaggerated the stringent shift in response bias from the first to the second showup. To the extent that participants thought that our procedure was similar to a police investigation, the short delays might have led participants to infer that there were many suspects available or that police were not investigating these suspects thoroughly. In either case, one would expect this to increase

the stringency of response bias. If actual investigations involve longer delays between showups, the criterion shift might be smaller, though this would mean that innocence risk would be even worse. Future research might consider examining the use of two or three showups with a single eyewitness and increasing the delays between showups.

Conclusion

To our knowledge, this is the first package of studies to examine the use of multiple showups with a single eyewitness for single perpetrator crimes. Together, these studies provide compelling evidence that the use of multiple showups is dangerous. Furthermore, given the apparent inability of multiple showups with the same eyewitness to provide evidence of guilt, their use is primarily – and we would argue exclusively – of value for investigative purposes. Even as an investigative procedure the use of multiple showup procedures is limited due to how frequently multiple showups will mislead police officers. Given these limitations, multiple showups should only be used in exigent circumstances. Finally, defense lawyers would do well to question witnesses about potential use of showups, even if the prosecution has not raised the issue.

Chapter 4

Can Scarcity Instructions Reduce Innocent Suspect Identifications in Multiple-Showup Procedures?[7]

After crimes are committed, law enforcement personnel frequently request descriptions of persons and events from eyewitnesses. With a description of the perpetrator in hand, law enforcement personnel may search the vicinity of the crime for individuals fitting the description. If law enforcement personnel locate an individual fitting this description, they might decide to conduct a showup identification procedure. A showup is a procedure in which a single individual (the suspect) is presented to an eyewitness, to test law enforcement personnel's hypothesis that the suspect perpetrated some crime.

Showups serve an important role in the criminal justice system as they are used in instances in which lineups (a more reliable identification procedure; Steblay, Dysart, Fulero, & Lindsay, 2003; Wetmore, Neuschatz, Gronlund, Wooten, Goodsell, & Carlson, 2015) are not practical. Specifically, it is intended that showups be used when an individual becomes a suspect due to fitting a general description. Match-to-general-description is not probable cause for arrest (*Washington v. Lambert,* 1996; *R. v. Mann*, 2004), but is cause for a brief investigatory detention. Under these parameters, law

[7] This manuscript is under review by the Journal of Police and Criminal Psychology:
Smith, A. M., Lindsay, R. C. L., & Myerson, T. (Under Review). Can scarcity instructions reduce innocent suspect identifications in multiple-showup procedures? *Journal of Police and Criminal Psychology.*

enforcement personnel would lack the time required to construct a lineup. The time needed to construct a lineup would require placing the suspect under arrest (Wells & Wilford, 2013).

Although showups serve an important role in criminal investigations, there is no denying that they can produce very high innocent suspect identification rates (e.g., Meissner, Tredoux, Parker, & MacLin, 2005, Experiment 4; Smith, Leach, & Cutler, 2013). Unlike lineups, in which a single suspect is embedded among known innocents (fillers hereafter), showups only include the lone suspect. Accordingly, when the perpetrator is not present, mistaken identifications from lineups are spread among lineup members (the innocent suspect is at less risk), but in showups, all mistaken identifications are innocent suspect identifications.

Worse yet, because showups are conducted in the field, law enforcement personnel have less control over their environment and showups may be host to variables that further increase the likelihood of mistaken identification. One such variable that has received considerable interest in the eyewitness literature is clothing bias. Past research finds that innocent suspect identifications are particularly likely when a suspect is presented in the same or highly similar clothing to that worn by the perpetrator (Dysart, Lindsay, & Dupuis, 2006; Wetmore, Neuschatz, Gronlund, Key, & Goodsell, 2015; Yarmey, Yarmey, & Yarmey, 1996).

Multiple-Showup Procedures

In addition to contextual variables present in a showup procedure, practices used by law enforcement might also increase innocent suspect identifications. One such practice is the use of multiple-showup procedures, the topic of the current paper. Smith, Bertrand, Lindsay, Kalmet, Grossman, and Provenzano (2014) asked law enforcement personnel if, after an eyewitness rejected a suspect they had gone out and found another suspect and conducted another showup with the same eyewitness. Many officers reported that they had used this procedure within a calendar year of participating in the study and some officers reported going through several iterations of finding suspects and running showups with the same eyewitness. In fact, nearly one fifth of law enforcement personnel who reported using a multiple-showup procedure within the previous calendar year, indicated that the maximum number of showups they had used with the same eyewitness was four or more. Anecdotally, in one of the most important pieces of American case law as it relates to eyewitness evidence, the eyewitness was presented with 30 to 40 photographs of suspects before she identified a suspect (*Neil v. Biggers,* 1972).

Multiple showup procedures are problematic because they provide eyewitnesses with more opportunities to identify innocent suspects, increasing the probability of an innocent suspect identification (Smith et al., 2014, Studies 2 – 4). Similar to multiple-suspect lineups, multiple-showup procedures do not put any *given* innocent at greater risk of mistaken identification, but they do increase the overall probability that *some* innocent

will be incorrectly identified (e.g., lineupwise error, Wells & Turtle, 1986). In fact, with as few as two showups, the probability that a suspect was innocent given identification (innocence risk, Clark & Godfrey, 2009) was greater than the probability that a suspect was guilty given identification. Warning participants (1) that the perpetrator may-or-may-not-be-present, (2) that clearing the innocent is just as important as implicating the guilty, and (3) that police will continue to investigate regardless of whether or not the eyewitness makes an identification, did decrease innocent suspect identifications; however, innocent suspect identifications were still unreasonably high due to the fact that they cumulate as the number of showups increase (Smith et al., 2014).

The present study had two primary objectives: (1) to examine whether slight modifications to these instructions could lead to more palatable levels of innocence risk in a multiple-showups procedure, and (2) to examine the contributions of two individual components of these instructions, base-rate information and perceptions of scarcity. We specifically examined the innocence risk in the instance that law enforcement personnel were willing to use four showups with the same eyewitness as this seemed to be a common upper bound to the number of showups law enforcement officers used with the same eyewitness in single-perpetrator crimes. Our primary focus was on the reduction of innocent suspect identifications for two reasons. First, as we have pointed out above, the probability of an innocent suspect identification in showups can be high for a number of reasons including (1) the absence of fillers, (2) the presence of contextual variables, and (3) the cumulative nature of false identifications when law enforcement personnel use

multiple showups. Second, there is good reason to suspect that the base-rate of target presence in showups is comparatively low when multiple-showup procedures are employed.

Showups Have a Comparatively Low Base-Rate of Perpetrator Presence

Showups are used when law enforcement personnel lack sufficient evidence to make an arrest, such as when an individual becomes a suspect based solely on fitting a general description of the perpetrator. In such instances, law enforcement personnel may put the suspect under investigatory detention and run a showup. In probabilistic terms, investigatory detention and showups are used when the prior probability that the suspect is the culprit (i.e., the *evidence* before an identification decision) is not sufficiently high to warrant an arrest. Thus, one might expect showups to include a high proportion of innocent suspects relative to identification procedures used when there is greater evidence that the suspect is the culprit (i.e., lineups). In the context of multiple-showup procedures, this is even more worrisome given that innocent suspect identifications cumulate. Furthermore, the average prior probability of perpetrator presence can only slightly exceed 25% and only to the extent that (1) the prior probability that the perpetrator would appear in the set of showups was high, and (2) information was gained from rejections (i.e., rejections decreased the probability that the suspect was guilty; e.g., Wells & Lindsay, 1980) on early showups.[8] It is apparent that showup procedures should be made as conservative as possible to safeguard against innocent suspect identifications.

[8] For a discussion of sequential updating in Bayesian analysis, see Lee & Wagenmakers (2013).

Identification Procedure Instructions

A large body of research suggests that the use of unbiased identification instructions decrease innocent suspect identifications and increase the reliability of identification (Steblay, 1997, 2013). That is, admonishing eyewitnesses that the perpetrator may-or-may-not be present in the lineup leads to a reduction in innocent suspect identifications. Given the practical benefits of the instruction, the Technical Working Group for Eyewitness Evidence (TWGEYE, 1999) recommended that before conducting an identification procedure, law enforcement personnel admonish all eyewitnesses that the perpetrator may-or-may-not be present. In fact, the TWGEYE (1999) went even further and recommended that law enforcement personnel also instruct eyewitnesses (1) that it is just as important to clear the innocent as it is to identify the guilty, and (2) that regardless of whether the eyewitness makes an identification, police will continue to investigate the incident.

Base-Rate Information

The may-or-may-not admonishment can be conceptualized as a manipulation of base-rate information. Although may-or-may-not instructions do not inform eyewitnesses of the probability that the culprit is present in the identification procedure, they do express that there is some uncertainty as to whether the perpetrator is present in the identification procedure. It is well known that decision-makers often neglect base-rate information (e.g., Kahneman & Tversky, 1973); however, individuals will use base-rate information if they believe that this information is responsible for the outcome

(Ajzen, 1977). If an eyewitness views an identification procedure and does not recognize anyone, the eyewitness might search for causal explanations to explain this recognition "failure". In the absence of a may-or-may-not admonishment, the eyewitness might not consider the possibility that the perpetrator is not present and may resort to an informed guessing strategy (e.g., relative judgments, Wells, 1984). The may-or-may-not admonishment affords the eyewitness another explanation for his or her failure to recognize the perpetrator – the perpetrator is not present (see Steblay, 2013; Charman & Wells, 2007 for similar derivations).

Based on this reasoning, we hypothesized that a more explicit base-rate information instruction might reduce innocent suspect identifications even further. We provided all participants in the present study with the may-or-may-not admonition. In addition, we instructed half of our participants that police do not always find the correct person. We reasoned that such an unexpected – and perhaps shocking – statement might be more salient to eyewitnesses who do not immediately recognize someone in an identification procedure and therefore, might further decrease innocent suspect identifications.

Perceptions of Scarcity

In addition to manipulating base-rate information, we also manipulated participants' perceptions of scarcity. As we have noted above, the TWGEYE (1999) recommended, in addition to the may-or-may-not admonishment, that law enforcement personnel also instruct the eyewitness that law enforcement personnel will continue to

investigate the incident regardless of whether or not the eyewitness makes an identification. Such an instruction might be construed as an operationalization of scarcity. Essentially, the instruction is intended to disabuse the eyewitness of any notion that opportunities to identify the perpetrator are scarce. Given that these instructions were recommended by the TWGEYE (1999), we were surprised to find that no research has investigated their utility.

Although we do not know of any research that has manipulated the "continue-to-investigate" instruction, there is a body of research in the eyewitness literature that might be interpreted as a manipulation of scarcity, namely, backloading. The sequential lineup is a procedure in which (1) lineup members are presented one at a time, (2) the eyewitness must make a decision about each member before proceeding to the next, (3) the eyewitness cannot return to previously rejected members, and (4) the eyewitness does not know how many members the lineup contains (i.e., the lineup is backloaded, Lindsay & Wells, 1985). Backloading is the practice of not telling eyewitnesses how many members the lineup includes or leading eyewitnesses – implicitly or explicitly – to believe that the lineup includes more members than it actually does. While the first three components of the sequential lineup were intended to limit intra-lineup comparisons (i.e., relative judgments, Wells, 1984), backloading was intended to prevent an eyewitness from reducing his or her decision criterion as the end of a lineup approached.

An eyewitness who is approaching the known end of a lineup might perceive opportunities to identify the perpetrator as becoming increasingly scarce and reduce his

or her decision criterion in order to seize the opportunity to identify the perpetrator. But, if the eyewitness does not know how many members the lineup includes or believes the lineup includes more members than it does (i.e., if the lineup is backloaded), then the eyewitness may maintain his or her decision criterion as the end of the lineup approaches. We are aware of two studies that have manipulated backloading (Horry, Palmer, & Brewer, 2012; Lindsay, Lea, & Fulford, 1991, Experiment 3). Lindsay et al. (1991) presented participants with a simultaneous lineup, a nonbackloaded-sequential lineup, or a backloaded-sequential lineup. All lineups included six members and all participants were admonished that the perpetrator may-or-may-not be present. In the backloaded-sequential lineup, participants were not told how many members the lineup included; however, the response form included 12-spaces. Innocent suspect identifications occurred more frequently in the nonbackloaded-sequential lineup than in the backloaded-sequential lineup; though both sequential lineups produced fewer innocent suspect identifications than the simultaneous lineup (Lindsay et al., 1991).

Horry et al. (2012) more directly addressed the mechanism through which backloading decreases innocent suspect identifications by manipulating target-presence, suspect position, and backloading. Horry et al. (2012) explicitly informed participants that they would see more lineup members than they actually would. Backloading reduced the identification of late-positioned innocent suspects and eliminated a tendency for eyewitnesses to gradually shift response bias as the end of the lineup approached (Horry et al., 2012). This pattern of results fits nicely with a scarcity explanation. When

an eyewitness knows that the end of a lineup is approaching, he or she perceives that opportunities to identify the perpetrator are becoming scarce and the eyewitness adjusts his or her response bias accordingly. Of course, if an eyewitness believes that opportunities to identify the perpetrator remain plentiful – as would be the case in the presence of backloading – then the eyewitness has no reason to adjust his or her response bias.

Both Lindsay et al. (1991) and Horry et al. (2012) found that backloading was effective at decreasing innocent suspect identifications in the presence of the may-or-may-not admonition. This suggests that the may-or-may-not admonition and the continue-to-investigate instruction might work additively. Indeed we hypothesized that decreasing eyewitnesses perceptions of scarcity would decrease innocent-suspect identifications over and above the may-or-may-not admonition. Instead of implying that eyewitnesses might have additional opportunities to identify the perpetrator – as the continue-to-investigate instruction does – however, we guaranteed the eyewitness additional opportunities to identify the perpetrator. Indeed, this practice is already in use by at least one law enforcement agency in the United States. All officers with the Norwood Police Department are provided with instruction cards to carry while on duty. On one side of the card, officers are provided with the Miranda Warnings. On the other side of the card, officers are provided with instructions to provide to the eyewitness when using a showup procedure. The very first instruction is to tell eyewitnesses that they are

going to be asked to view *some* people, even if only one person is shown (William G. Brooks,[9] personal communication, March 20, 2015).

In summary, given that showups are comparatively likely to have low prior probabilities of perpetrator presence and because multiple-showup procedures lead to high rates of innocent suspect identifications, we examined instructions designed to reduce innocent suspect identifications. Specifically, we examined whether instructing eyewitnesses that the police do not always get the correct person could decrease innocent suspect identifications over and above the standard may-or-may-not be present admonition. Additionally, we examined whether decreasing perceptions of scarcity might also decrease innocent suspect identifications, as is the case with backloading in sequential lineups. Finally, we examined the impact of these instructions on the likelihood that an eyewitness would survive (i.e., reject all innocents) to see an eyewitness in the fourth showup and how likely these eyewitnesses were to identify the perpetrator in the fourth showup.

Method

Participants and Design

Participants ($N = 231$) from a university in Southern Ontario participated individually in exchange for course credit. We recruited participants through the psychology subject pool and through campus e-mail. Participants on average were 18.26 ($SD = 0.76$) years of age and 86.1% of participants were female. The majority of

[9] Chief of police, Norwood Police Department, Norwood, Massachusetts, USA. Copies of showup instructions and other materials can be obtained from: www.norwoodpolice.com.

participants were of European ancestry: European (67.1%), Asian (21.6%), African (2.2%), Middle Eastern (1.3%), First Nations (1.3%) and "Other" (6.5%).

All participants received a may-or-may-not admonition. The first three showups were always target-absent and the fourth was always target-present. Participants were randomly assigned to a 2 (additional base-rate instruction: absent, present) × 2 (scarcity instruction: absent, present) between-subjects design.

Materials

Targets

Each participant was presented with only a single target video; however, we used three different targets in this study to attain some limited degree of stimulus sampling (Wells & Windschitl, 1999). Each target was yoked to three innocent suspects who each fit the general description of their respective target. All targets and respective innocents were male, European, in their early 20s, and varied in hair length and colour (very short, light brown hair; medium-length, very dark hair; medium-length blonde hair).[10] None of our outcome measures varied as a function of stimuli; therefore, we do no differentiate among them in our analyses.

Each target video lasted approximately 2 s in duration, depicted the target's face from a three-quarter angle, and included a brief statement from the target related to criminal activity (e.g., "So, there's good stuff in that place?"). Each video was approximately 18 cm × 25 cm. In each showup photo, the face of the suspect was

[10] These stimuli are the same stimuli that were used in Smith et al. (2014); for average similarity ratings see Smith et al. (2014).

depicted from the neck up (no clothing cues) looking directly at the camera. Each photograph was approximately 22 cm x 28 cm.

Filler tasks

After the encoding task, participants completed an unrelated study that lasted approximately 30 minutes. After the delay and in between showups, participants solved anagrams by rearranging scrambled letters to create words. The duration of the anagrams task varied from 30s to 194s between showups in order to limit the ability of participants to anticipate when (or if) another showup might occur, but the inter-trial intervals were the same for every participant.

Procedure

Participants were brought into a room, individually, and instructed to watch a video on a computer. Following the video, participants read a letter of information on screen, gave consent, and provided demographic information. After providing demographic information, participants completed an unrelated study that lasted approximately 30 minutes. This delay was included to increase the memory-retention interval.

Following completion of the unrelated study, participants completed the anagrams filler task. The task was interrupted by a set of instructions appearing on screen regarding the identification procedure. The control condition received only the may-or-may-not be present admonishment. Participants receiving additional base-rate information were instructed that “Police do not always find the right person. For that

reason, the person you will be shown may or may not be the man from the video." Participants receiving scarcity information were instructed that "If you do not believe that the guy in the photo is the man you saw, others will be shown later for you to attempt to identify the guy from the video." After the participant was provided with the instructions, a face appeared on screen and participants were asked if they believed this was the man from the video. Participants were forced to provide a definitive "Yes" or "No" response. Participants who responded "No" returned to the anagrams task. After a predetermined amount of time had elapsed, participants were presented with another face and asked if it was the man from the video. The procedure continued until the participant either made an identification or rejected all four showups.

Results

Correct Rejections

We first examined the impact of the instructions on the probability that eyewitnesses would reject the first innocent suspect. Because we were interested in the main effects of instructions and the interaction term, we used a hierarchical logistic regression model. We entered main effects on the first block and the interaction term on the second. We used this same procedure for all subsequent logistic regression models. Participants who received a scarcity instruction were significantly more likely to correctly reject the first innocent suspect (P = .90, SD = .31)[11] than were those participants who did not receive a scarcity instruction (P = .78, SD = .42), B = .91, SE =

[11] P for proportion as is recommended by Cohen, Cohen, West, and Aiken (2003).

.38, Wald's $\chi^2(1) = 5.84$, $p = .02$, $e^B = 2.50$ (95% CI [1.19, 5.24]). Participants who were provided with the scarcity instruction (8.58:1) were 2.50 times as likely to identify the first innocent suspect, than were participants who did not receive the scarcity instruction (3.46:1).[12] Neither the additional base-rate instruction, nor the interaction between additional base-rate and scarcity instructions significantly influenced the likelihood of a correct rejection, Bs $\leq .46$, SEs $\geq .36$, Wald's $\chi^2(1) \leq 0.61$, ps $\geq .44$, e^Bs ≤ 1.56. Proportions of suspect choices for all showups are reported in Table 7. Note that the suspect identifications reported in the table provide the same information as the correct rejections discussed in text, as the proportion of innocent suspect identifications is equal to one minus the proportion of correct rejections.

[12] The values in parentheses represent the odds of an innocent suspect identification for a given condition. To find the odds ratio, one divides the odds for one condition by the odds for the condition to which it is being compared.

Table 7: *Frequency of Suspect Selections by Instructions and Showup*

	1	2	3	4	None
Instructions					
None (*n*= 57)	15 (26%)	9 (16%/**21%**)	5 (9%/**15%**)	7 (12%/**25%**)	21
Base-Rate (*n*= 59)	11 (19%)	14 (24%/**29%**)	11 (19%/**32%**)	6 (10%/**26%**)	17
Scarcity (*n*= 58)	6 (10%)	12 (21%/**23%**)	7 (12%/**18%**)	13 (22%/**39%**)	20
Both (*n*= 57)	6 (11%)	13 (23%/**26%**)	5 (9%/**13%**)	13 (23%/**39%**)	20
Totals	38	48	28	39	78

Note. The first percentage is the unconditional proportion of suspect selections, while the second is the proportion of suspect selections conditional on the participant surviving to see the suspect. Target responses are underlined. The no instruction condition did include the may-or-may-not admonition, but this instruction was provided to all participants.

Next, we examined how many correct rejections participants made, on average, as a function of the pre-identification instructions. Given that the first three showups contained innocent suspects, participants could make a maximum of three correct rejections (range: 0 – 3). Participants who received scarcity instructions, on average, correctly rejected more innocent suspects ($M = 2.15$, $SD = 1.09$) than those who did not receive scarcity instructions ($M = 1.79$, $SD = 1.23$), $F(1, 227) = 5.32$, $p = .02$, $d = 0.33$. Neither the additional base-rate instructions nor the interaction term were significant, Fs$(1, 227) \leq 0.03$, ps $\geq .89$. See Table 8 for the average number of correct rejections for each condition.

We then examined the likelihood that participants would survive to see the target in the fourth showup as a function of the pre-identification instructions. Participants who received scarcity instructions were significantly more likely to survive to see the target in the fourth showup (P = .57, SD = .49) than were participants who did not receive scarcity instructions (P = .44, SD = .50), B = .54, SE = .27, Wald's $\chi^2(1) = 4.11$, p = .04, e^B = 1.72 (95% CI [1.02, 2.89]). Participants who were provided with the scarcity instruction (1.35:1) were 1.72 times more likely to survive to see the target in the fourth showup, than were participants who did not receive the scarcity instruction (0.78:1). Neither the additional base-rate instruction, nor the interaction between the additional base-rate instruction and the scarcity instruction increased the likelihood of surviving to see the target, Bs = -.19 to .45, SEs ≥ .27, Wald's $\chi^2(1) \leq 0.73$, ps ≥ .39, e^Bs = .83 to 1.57. See Table 7.

Table 8: *Impact of Instructions on Average Number of Correctly Rejected Innocent Suspects*

	n	*M*	*SD*	95% CI for Mean	
				Lower Bound	Upper Bound
None	57	1.81	1.30	1.50	2.11
Base-Rate	59	1.78	1.16	1.48	2.08
Scarcity	58	2.16	1.09	1.85	2.46
Base-Rate & Scarcity	57	2.14	1.11	1.84	2.46

Note. Maximum correct rejections possible was three, as the fourth showup was always target-present. Confidence intervals are descriptive, not inferential. The no instruction condition did include the may-or-may-not admonition, but this instruction was provided to all participants.

Hits

We also examined the likelihood that participants would identify the perpetrator as a function of the identification instructions. Participants who received scarcity instructions were significantly more likely to identify the target in the fourth showup (P = .23, SD = .42), than were participants who did not receive scarcity instructions (P = .11, SD = .32), B = .84, SE = .37, Wald's $\chi^2(1) = 5.16$, p = .02, e^B = 2.31 (95% CI [1.12, 4.77]). Participants who were provided with the scarcity instruction (0.29:1) were 2.31 times more likely to survive to identify the target in the fourth showup, than were participants who did not receive the scarcity instruction (0.13:1). Neither the additional base-rate instruction, nor the interaction between additional base-rate and scarcity instructions increased the likelihood of a correct identification, Bs = -.06 to .24, SEs ≥ .36, Wald's $\chi^2(1) \leq 0.10$, ps ≥ .75, e^Bs = 0.94 to 1.27.

Of course, more participants who received the scarcity instruction survived to see the target in the fourth showup. Accordingly, we also examined the proportion of participants who identified the target, given that they survived to see the target. The conditional probability of a target identification did not vary as a function of the pre-identification instructions, Bs = .02 to .64, SEs ≥ .40, Wald's $\chi^2(1) \leq 2.46$, ps ≥ .12, e^Bs = .944 to 1.90. See Table 7.

Innocence Risk

Finally, we also examined innocence risk (Clark & Godfrey, 2009) as a function of the pre-showup instructions. Innocence risk is the probability that a suspect is

innocent given that the suspect was identified. Following Smith et al. (2014), our measure of cumulative risk reflected the fact that innocent suspect identifications cumulate and that not all eyewitnesses will survive to see the target in the fourth showup. Hence, our false alarm rate reflected the probability that an eyewitness would identify one of the first three innocent suspects and our hit rate was the unconditional proportion of participants who identified the perpetrator. Because we wanted to inferentially compare the innocence risk associated with the different instructions, we used a modified jackknife procedure (Mosteller & Tukey, 1968; Smith et al., 2014) to estimate the standard errors associated with the given innocence risk values. We then used the estimated standard errors to calculate inferential 95% confidence intervals around each statistic (Tryon, 2001; see Table 9).. Innocence risk was significantly lower when participants were provided with the scarcity instruction (innocence risk = .66, *SE* = .07, 95% CI [.57, .75]) or both the scarcity instruction and the additional base-rate instruction (innocence risk = .65, *SE* = .07, 95% CI [.55, .74]) than when participants were provided with only the additional base-rate instruction (innocence risk = .86, *SE* = .05, 95% CI [.79, .93]). Those who received scarcity instructions also trended towards lower levels of innocence risk when compared to those who received neither instruction. See Table 9. Yet, innocence risk exceeded .50 for all instructional conditions, meaning that identified suspects were always more likely to be innocent than guilty.

Table 9: *Innocence Risk as a Function of the Pre-Showup Instructions*

		Jackknife Statistics	
Instructions	Innocence Risk	*SE*	95% CI
None	.81	.06	[.72, .89]
Base-Rate	.86	.05	[.79, .93]
Scarcity	.66	.07	[.57, .75]
Both	.65	.07	[.55, .74]

Note. We do not report the mean of the jackknife procedure as in all instances it was equal to the descriptive innocence risk value. The no instruction condition did include the may-or-may-not admonition, but this instruction was provided to all participants.

Discussion

Our attempt to increase the robustness of the may-or-may-not admonition was unsuccessful; however, instructing participants that they would have additional opportunities to identify the perpetrator significantly decreased innocent suspect identifications on the initial showup and increased the proportion of participants who survived to see the perpetrator in the fourth showup. This also would have been true if the perpetrator had been presented in the second or third showup. Decreasing participants' perceptions that opportunities to identify the perpetrator were scarce effectively decreased innocent suspect identifications over and above the may-or-may-not admonition. Although results for the scarcity instructions were promising, innocence risk was high no matter the instructions that were provided before the showup identification procedure.

Theoretically, we believe these results demonstrate that scarcity is an important mechanism in determining an eyewitnesses' willingness to choose. Since the inception of the sequential lineup, researchers have acknowledged that eyewitnesses might become more willing to choose as the end of the lineup approached (Lindsay & Wells, 1985). Indeed, research on backloading in sequential lineups has demonstrated that, in the absence of backloading, eyewitnesses decrease decision criteria and thus increase choosing as the end of the lineup approaches (Horry et al., 2012). Presumably this increased choosing reflects the fact that eyewitnesses recognize that few opportunities to identify the perpetrator remain. Here we have demonstrated that perceiving opportunities to identify the perpetrator as scarce also contributes to an increased willingness to choose in showup identification procedures. The scarcity mechanism is likely to be an important consideration in all identification tasks.

From an applied perspective, multiple-showup procedures proved to be extremely dangerous. Even in the presence of relatively ideal conditions – a may-or-may-not admonition, an additional base-rate instruction, and a scarcity instruction – innocence risk was unacceptably high. It seems safe to conclude from our results and those of Smith et al. (2014) that a willingness to use four showups with the same eyewitness is never going to lead to solid evidence of guilt. Due to the fact that innocent suspect identifications cumulate, even the most conservative of instructions are insufficient to produce palatable levels of innocence risk (innocence risk values <.50). If identified suspects are more likely to be innocent than guilty from some procedure, it is hard to argue that this

procedure is useful either as an investigative tool or as a source of courtroom evidence. This procedure will simply send police investigations awry far too often to rely on it as a powerful investigative tool. Even worse, innocence risk values consistently above .5 indicate that an identification from a showup, and particularly repeated showups, should never be used as evidence of guilt in court. Police certainly could use a series of showups to find a suspect worthy of further investigation; however, the showup identification will always lack probative value, though the suspect could be viewed in a lineup by other eyewitnesses who had not been involved in the showup procedure.

On a more positive note, we did find that informing participants that they would have additional opportunities to identify the perpetrator significantly reduced innocent suspect identifications. In the present study, participants were actually given additional opportunities to identify the perpetrator as we examined a multiple-showups procedure; however, the instruction would have equal utility in a single-showup or –lineup procedure. Indeed, one can simply think of this instruction as backloading as does the Norwood Police Department. Given that the continue-to-investigate instruction was recommended by the TWGEYE (1999), we were surprised that it had received no empirical attention. It seems to be a promising avenue for future research as it decreased innocent suspect identifications over and above the may-or-may-not admonition.

There are a couple of notable limitations to the present research. First, we underestimated the innocence risk associated with a willingness to use four showups with the same eyewitness as we did not have a condition in which participants could see a

fourth innocent suspect. This could not affect our conclusion that multiple showup procedures are dangerous as having this additional target-absent showup could only increase the cumulative false identification rate and the resultant innocence risk.

We also did not have conditions in which participants could see the target in the first, second, or third showups as our primary concern was with reducing innocent suspect identifications. Accordingly, it is possible that the scarcity instruction might decrease perpetrator identifications to the extent that it cancels out any benefit stemming from reducing innocent suspect identifications. Future research should examine the full 2 × 2 orthogonal design in single-identification procedures to establish that any reduction in innocent suspect identifications is not completely offset by a reduction in perpetrator identifications.

Conclusion

To our knowledge, this is the first study to systematically manipulate the presence or absence of scarcity instructions in identification procedures. This study provides evidence that an instruction intended to decrease perceptions of scarcity might have utility in reducing innocent suspect identifications over and above the may-or-may-not admonition. In addition, this is only the second paper we are aware of to examine the use of multiple-showup procedures. Buttressing the findings of Smith et al. (2014), our results suggest that the use of multiple showup procedures lead to unacceptably high levels of innocence risk. We would go so far as to suggest that our results, coupled with those of Smith et al. (2014), suggest that when law enforcement personnel are willing to

use multiple showups with the same eyewitness, identifications are *never* evidence of guilt.

Chapter 5

Conclusion

In summary, this dissertation demonstrates the dangers associated with using multiple-showup procedures. As the number of showups law enforcement are willing to use with the same eyewitness increases, so too does the probability that a suspect is innocent given identification (innocence risk, Clark & Godfrey, 2009). Multiple-showup procedures do not increase the risk of mistaken identification for any given innocent suspect, but they do increase the risk that some innocent person will be identified, because innocent suspect identifications cumulate. To find the cumulative innocent suspect identification rate associated with a multiple-showup procedure, it is necessary to sum together the probabilities of an innocent suspect identification on each of the individual showups. For example, if law enforcement personnel are willing to use three showups with the same eyewitness, the cumulative innocent suspect identification rate is found by summing the probabilities of an innocent suspect identification in the first, second, and third showups. Making matters worse, if the perpetrator is not in the initial showup, the probability of a perpetrator identification will decrease as many eyewitnesses will choose innocent suspects before ever seeing the perpetrator. While I did find some evidence that pre-showup instructions can decrease the risk of innocent suspect identification, innocence risk remained unacceptably high (Chapter 3: Study 4, and Chapter 4).

This dissertation also demonstrates the importance of considering eyewitness' perceptions of scarcity. Instructing participants that they would have additional opportunities to identify the perpetrator decreased innocent suspect identifications over and above the use of a pre-showup admonition indicating that the perpetrator may-or-may-not be presented. Perceptions of scarcity may be an important mechanism to consider in identification procedures more generally.

In this concluding chapter, I argue that identifications from multiple-showup procedures are not reasonable evidence of guilt. Given the high rates of innocence risk associated with multiple showup procedures, the prosecution must disclose all identification procedures, no matter the response of the eyewitness. Next, I examine the impact of pre-showup instructions and the importance of ensuring that eyewitnesses do not perceive opportunities to identify the perpetrator as scarce. Furthermore, I suggest that scarcity is an important mechanism to consider in identification procedures more generally. I then review the limitations of this program of research and raise questions to be addressed in future research. Finally, I conclude with discussion of a need for a stronger partition between investigative and evidentiary procedures.

Identifications from Multiple-Showup Procedures are Not Reasonable Evidence of Guilt

Given the overwhelming probability that a suspect is innocent given identification, it seems reasonable to argue that an identification from a multiple-showup procedure should never be considered sufficient evidence for conviction. Study 3 in

Chapters 3 demonstrated that innocence risk was in excess of 50% with as few as two showups. In other words, these data suggest that, when law enforcement personnel are willing to use two or more showups with the same eyewitness, identified suspects are more likely to be innocent than guilty. Granted, I only examined the innocence risk associated with two showups in one study (Chapter 3, Study 3) and I did not provide participants with pre-showup instructions in this study. In the presence of pre-showup instructions, innocence risk might be less than 50%; however, this possibility should not be taken as a victory. Arguing that a procedure might have efficacy if it produces innocence risk values less than 50% is using the most lax criterion that could sensibly suggest a procedure might be useful. Indeed, an innocence risk value of 50% would suggest that the suspect is equally likely to be innocent or guilty and values greater than 50% suggest that the suspect is more likely to be innocent than guilty. Even if a pre-showup admonition decreased innocence risk in a two-showup procedure to 40% for example, identifications would still only be weak evidence of guilt as 40% of identifications would be of innocent suspects.

Furthermore, the probability that one is innocent given identification is even more concerning after considering the base-rate of target presence. As argued in Chapter 4, the prior probability (i.e., the base rate) that a suspect in a showup procedure is the perpetrator is likely lower than in a lineup procedure. Showups are often used when law enforcement personnel lack sufficient evidence to make an arrest (i.e., when there is a low prior probability that the suspect is the culprit). In probabilistic terminology, law

enforcement personnel use showups when the prior probability that the suspect is the perpetrator is low.

Making matters worse, when multiple showup procedures are used, the probability that any given suspect is the culprit is likely to be exceedingly low. Mathematically, when law enforcement personnel are willing to use two showups with the same eyewitness for a single-perpetrator crime, the average prior probability can only slightly exceed .50. Before the eyewitness makes any identification decision, the sum of the priors must fall between 0, *a priori* certainty that the suspect will not appear in either showup, and 1, *a priori* certainty that the suspect will appear in one of the two showups. Hence, .50 would be the maximum average prior probability for a given showup. If the eyewitness rejects the initial showup, this will decrease the probability that the suspect is the culprit, because rejections on the initial showup were more common than identifications when the suspect was innocent. To the extent that a rejection on the initial showup decreases the probability that the suspect is the perpetrator, the sum of the prior probabilities for showups one and two may exceed 1.00 (see Lee & Wagenmakers, 2012 for a discussion of sequential updating in Bayesian analysis). The average prior probability in this instance would only exceed .50 to the extent that (1) the prior probability that the perpetrator would appear in one of the two showups was high to begin with, (2) information was gained from a rejection in the initial showup (i.e., a rejection in the initial showup decreased the probability that the suspect was the perpetrator; Wells & Lindsay, 1980), and (3) the information gained from a rejection on

the initial showup is added to the probability that the suspect in the second showup is the perpetrator. Even then, it is not clear why the information gained from rejecting a suspect on the initial showup would increase the probability that the suspect in the second showup was the perpetrator; but, in theory, it could.

Accordingly, when law enforcement personnel are willing to use two showups with the same eyewitness, the average prior probability that a suspect in a given showup is the perpetrator is not likely to greatly exceed .50. Conversely, the prior probabilities could be exceedingly low – there is nothing preventing a prior probability from being 0. What this suggests is that the innocence risk values presented in Chapters 3 and 4 are very close to the best-case scenario as those values are based on the implicit assumption that the prior probabilities are .50.[13] This value would be close to the upper limit for a two-showup procedure and would be a gross overestimation of the average prior in the instance that law enforcement personnel were willing to use three or four showups with the same eyewitness. This logic suggests that the use of multiple-showup procedures might be far more dangerous than my experiments suggest.

To the extent that the prior probability that the suspect is the perpetrator is low, the probability that the suspect is the perpetrator given identification (the posterior probability) will also be low (Wells & Lindsay, 1980; Wells & Turtle, 1986; Wells et al., 2015). That is to say, the base-rate of target presence is informative of the likelihood that

[13] When one uses measures of probative value (e.g., innocence risk, Clark & Godfrey, 2009) to compare procedures, it is rare to specify prior probabilities given that these values are largely unknown and vary considerably among law enforcement personnel (Wells, Yang, & Smalarz, 2015). When one does not specify prior probabilities with Bayesian measures, the individual is tacitly assuming a prior probability of .50.

an identified suspect is the perpetrator. Given that the use of multiple-showup procedures is informative of the prior probability that the suspect was the perpetrator (i.e., the base-rate) and the prior probability informs as to the likelihood that a suspect is innocent given identification, all identification procedures, no matter the decision made by the eyewitness, *must* be disclosed to the defense.

In both Canada and the United States, prosecutors must disclose all potentially exculpatory information to the defense (*Brady v. Maryland,* 1963; *R. v. Stinchcombe,* 1991). Of course, it is unlikely that individuals – prosecutors or otherwise – would recognize that the number of showups law enforcement personnel used with an eyewitness is potentially exculpatory. Although it seems unlikely to me that the number of showups an eyewitness rejected would be entered into evidence, it seems probable that if it were, it would be argued as evidence in support of the prosecution's case. In reality, one could envision a prosecutor discussing the number of rejections an eyewitness made before identifying the suspect to support the reliability of the eyewitness. For example, in the case of *Neil v. Biggers* (1972), Neil might have come across as one of the most reliable eyewitnesses many officers had ever seen given that she rejected some 40 suspects before identifying Biggers as her assailant. But, consider the base-rate of target presence in the instance that law enforcement personnel present 40 different suspects to the same eyewitness. Even if finding that many suspects could guarantee that the perpetrator was among them, the average prior probability for each given suspect would

still only be about .025 (i.e., on average, there would be a 2.5% chance that any given suspect was the perpetrator before identification).

What should be clear from the above discussion is that when law enforcement personnel use multiple showups with the same eyewitness, this provides important information about the probability that a suspect was the culprit before an identification. Because the probability that a suspect was the culprit before an identification constrains the probability that a suspect is the perpetrator given identification, knowledge that law enforcement personnel use multiple-showup procedures with the same eyewitness is relevant to evaluating a given case and should be disclosed to the defense. Given the high risk of innocent suspect identification and the low prior probabilities associated with multiple-showup procedures, it seems reasonable to suggest that an individual should never be convicted solely on the bases of an identification from a multiple-showup procedure. Going one step further, defense lawyers might argue that the identification evidence from a multiple-showup procedure is not probative and should not be admissible as evidence at all.

The Influence of Pre-Showup Instructions on the Reliability of Identification

Providing eyewitnesses with the instructions recommended by the Technical Working Group for Eyewitness Evidence (TWGEYE, 1999) decreased innocent suspect identifications; however, the probability that a suspect was innocent given identification remained unacceptably high (Chapter 3, Study 4). Specifically, the TWGEYE (1999)

recommended that law enforcement personnel instruct eyewitnesses that (1) the perpetrator may-or-may-not be present, (2) it is just as important to reject the innocent as it is to identify the guilty, and (3) law enforcement personnel will continue to investigate whether the eyewitness makes an identification or not. In Chapter 3, Study 4, I presented these instructions as a package, as opposed to manipulating those components participants received.

In Chapter 4, I sought to examine whether more explicit manipulations of base-rate information or perceptions of scarcity might reduce innocence risk to more palatable levels. Unfortunately, innocence risk remained unacceptably high. Yet, I did find one positive result. Telling eyewitnesses that they would have additional opportunities to identify the perpetrator if they did not make an identification decreased innocent suspect identifications over and above a may-or-may-not be present admonition. This is a very important finding. Although the TWGEYE (1999) recommends that law enforcement personnel provide eyewitnesses with a "continue-to-investigate" instruction, there is no research to my knowledge that has examined the efficacy of this instruction. The study presented in Chapter 4 was the first to demonstrate that instructions can decrease perceptions of scarcity and that this can lead to a decrease in innocent suspect identifications over and above a may-or-may-not be present admonition. This finding has both applied and theoretical implications.

At the applied level, finding that decreasing participants perceptions of scarcity decreases innocent suspect identifications suggests that this instruction should be used in

addition to the may-or-may-not be present admonition. Although the TWGEYE (1999) recommended a similar continue-to-investigate instruction, the manipulation of biased versus unbiased instructions generally boils down to the presence or absence of the may-or-may-not be present admonition (Steblay, 2013). But, the research presented in Chapter 4 suggests that researchers should not dismiss the efficacy of scarcity instructions. In showup identification procedures in particular, where the prior probability that the suspect is the perpetrator is likely to be lower, telling eyewitnesses that they will have additional opportunities to identify the perpetrator may be particularly useful at decreasing innocent suspect identifications.

At the theoretical level, the research presented in Chapter 4 in concert with research on backloading (Horry, Palmer, & Brewer, 2012; Lindsay, Lea, & Fulford, 1991, Experiment 3) suggests that perceptions of scarcity might be an important mechanism to consider in all identification procedures. Indeed, both the scarcity instruction used in the present research and backloading have shown the ability to effectively decrease innocent suspect identifications. Presumably, both backloading and the scarcity instruction are effective because they increase the extent to which eyewitnesses perceive opportunities to identify the perpetrator as plentiful. The fact that scarcity instructions or backloading can reduce innocent suspect identifications suggests that, in the absence of some manipulation intended to decrease perceptions of scarcity, some eyewitnesses will make an identification out of fear that they will not have additional opportunities to do so. It is probably safe to assume that the legal system

would not find it desirable to obtain an identification if it is attributable to an eyewitness' fear that he or she will not have additional opportunities to identify the perpetrator. It follows that law enforcement personnel should disabuse eyewitnesses of such perceptions before any identification procedure.

Limitations and Directions for Future Research

There are two notable limitations to the research presented in this dissertation. First, as I have alluded to above, in the weakest sense of the word, a two-showup procedure may have some utility in the instance that the showups are preceded by pre-identification instructions. Yet, I have also pointed out above that in the instance that law enforcement personnel use multiple-showup procedures, the average prior probability that any given suspect is the perpetrator is likely to be exceedingly low. After considering the prior probabilities, it is unlikely that even a two-showup procedure with pre-showup instructions would have any probative value in the direction of guilt.[14] Nevertheless, future research would do well to examine the innocence risk associated with two showups when the showups are preceded by pre-showup instructions.

[14] From a legal standpoint, a piece of evidence is probative if it makes some fact more or less probable than that fact would be without the piece of evidence (*R. v. Watson,* [1996] O.J. No. 2695 at paragraph 33). Indeed, this is consistent with how eyewitness researchers think of probative value and as pointed out in Chapter 2 of this dissertation, it is often represented as a likelihood function called the diagnosticity ratio (Wells & Lindsay, 1980); a value of (or close to) 1 would indicate that a fact has no probative value. The point here is that, identifications from a two-showup procedure are limited in there ability to inculpate and even if procedures could increase the reliability of the identification decision, the constraints placed on the prior probabilities would attenuate this benefit (because the priors have probative value in the direction of innocence). In effect, the likely probative value associated with an identification from a two-showup procedure is likely to be nil or in the direction of innocence.

In order to examine innocence risk for two showups in the presence of pre-showup instructions, participants would need to be randomly assigned to view the perpetrator in the first, second, or neither showup. Randomly assigning some participants to view the perpetrator in the first showup is necessary to establish baseline performance. The other two conditions are necessary to calculate the innocence risk associated with two showups. The appropriate study could be an exact replication of Study 3 in Chapter 3 with the exception that all showups would be preceded by pre-showup instructions. It might be the case that under such conditions, innocence risk would fall below .50, meaning that identified suspects are more likely to be guilty than innocent. Yet, as argued above, it remains unlikely that such a procedure would be highly reliable as large proportions of identified suspects would still likely be innocents. Nevertheless, future research should empirically examine these predictions.

The second notable limitation is that I did not examine the impact of scarcity instructions on the rate of perpetrator identifications in the first showup. Scarcity instructions may be useful in showups and other identification procedures to the extent that perpetrator identifications do not plummet in their presence. Some researchers have argued that cautionary instructions merely lead to a shift in decision criterion and do not increase discrimination (Clark, 2012). Even still, decreasing innocent suspect identifications increases the reliability of identification decisions (Steblay, 2013). While it would be preferable for perpetrator identifications to remain intact in the presence of scarcity instructions, so long as innocent suspect identifications decrease to a

proportionally greater extent than perpetrator identifications, scarcity instructions will increase the reliability of identification. Future research should examine the utility of scarcity instructions on both innocent suspect and perpetrator identifications in single-suspect identification procedures.

Finally, it is important to consider how the utility of multiple showup procedures varies as a function of the prior probabilities. As noted above, the prior probability that the suspect in the second showup is the perpetrator is constrained by the prior probability that the suspect in the initial showup is the culprit. For example, if the prior probability that the suspect in the first showup is 0 (i.e., it is certain *a priori* that the suspect is not the perpetrator), then the prior probability that the suspect in the second showup is the perpetrator can range anywhere between 0 (*a priori* certainty that the suspect is not the perpetrator) and 1 (*a priori* certainty that the suspect is the perpetrator). Given that there is no chance that the suspect in the initial showup is the perpetrator, the prior probability that the suspect in the second showup is not constrained by the prior in the initial showup – the prior in the second showup can range from 0 to 1. If, however, the prior probability that the suspect in the first showup is the perpetrator is equal to 1 (i.e., if the first suspect in is the perpetrator), then the prior probability that the suspect in the second showup is the perpetrator must be 0 (i.e., the second suspect must be innocent). As the prior probability that the suspect in the initial showup is the perpetrator increases, the range of potential priors that the suspect in the second showup is the perpetrator decreases.

The development of Bayesian models that vary the prior probabilities associated with a given showup are important for determining when or if additional showups might have utility. In the instance that the prior probability in the first showup is low, a second showup may have utility, as the prior probability with that showup might be high. The question of import is how low must the prior in the first showup be for the second showup to have utility. In addition, such Bayesian models would permit the examination of how constrained priors and cumulating innocent suspect identification rates coalesce to produce exceedingly high innocence risk values.

Final Remarks

In summary, I do not believe that an identification from a multiple-showup procedure should ever be considered sufficient evidence to convict. It should be necessary that law enforcement find some converging evidence before convicting a suspect on the bases of an identification from a multiple-showups procedure. I am not suggesting that law enforcement personnel should not be able to use multiple-showup procedures with a single eyewitness. Although law enforcement personnel should use alternative investigative procedures when practicable, there will be instances in which law enforcement personnel have no leads other than the eyewitness. In such instances, it hardly seems reasonable to suggest that law enforcement personnel could not use multiple-showup procedures if the eyewitness rejected the first (and potentially subsequent) showup that law enforcement personnel presented. If law enforcement personnel do obtain an identification from a multiple-showup procedure, they should treat

the identification with a healthy dose of skepticism as the compendium of factors discussed in this dissertation suggest that there is a high probability that the suspect is innocent given identification. Additionally, while the use of multiple-showup procedures may prove necessary for investigative purposes, identifications from multiple-showup procedures should not be treated as evidence of guilt.

References

Agricola, B. H. (2009). The psychology of pretrial identification procedures: The show-up is showing out and undermining the criminal justice system. *Law and Psychology Review, 33,* 125 – 137.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35,* 303 – 314. doi: 10.1037/0022-3514.35.5.303

Andersen, S. M., Carlson, C. A., Carlson, M. A., & Gronlund, S. D. (2014). Individual differences predict eyewitness identification performance. *Personality and Individual Differences, 60,* 36 – 40. doi: 10.1016/j.paid.2013.12.011

Austin, J. L., Zimmerman, D. M., Rhead, L., & Kovera, M. (2013). Double-blind lineup administration: Effects of administrator knowledge on eyewitness decisions. In Cutler, B. L. (Ed.), *Reform of Eyewitness Identification Procedures.* (pp. 139 - 160). Washington, DC: American Psychological Association.

Beal, C. R., Schmitt, K. L., & Dekle, D. J. (1995). Eyewitness identification of children: Effects of absolute judgments, nonverbal response options, and event encoding. *Law and Human Behavior, 19,* 197 – 216. doi:10.1007/BF01499325

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behaviour, 25*, 475 – 491. doi: 10.1023/A:1012840831846

Behrman, B. W. & Vayder, L. T. (1994). The biasing influence of a police showup: Does the observation of a single suspect taint later identification? *Perceptual and Motor Skills, 79,* 1239 – 1248. doi:10.2466/pms.1994.79.3.1239

Brady v. Maryland (1963). 373 U.S. 83.

Buckhout, R. (1974). Eyewitness testimony. *Scientific American, 231*, 23 – 31.

Carlson, C. A. & Carlson, M. A. (2014). An evaluation of lineup presentation, weapon presence, and a distinctive feature using ROC analysis. *Journal of Applied Research in Memory and Cognition, 3*, 45 – 54. doi: 10.1016/j.jarmac.2014.03.004

Charman, S. D. & Wells, G. L. (2007). Applied lineup theory. In R. C. L. Lindsay, D.F. Ross, M. P. Toglia, & D. Read (Eds.), *The handbook of eyewitness psychology, volume 2: Memory for people.* (pp. 219 – 254). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology, 17,* 629 – 654. doi:10.1002/acp.891

Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior, 29,* 395 – 424. doi:10.1007/s10979-005-5690-7

Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science, 7,* 238 – 259. doi:10.1177/1745691612439584

Clark, S. E. & Godfrey, R. D. (2009). Eyewitness identification evidence and innocence risk. *Psychonomic Bulletin & Review, 16,* 22 – 42. doi:10.3758/PBR.16.1.22

Clark, S. E., Rush, R. A., & Moreland, M. B. (2013). Constructing the lineup: Law, reform, theory, and data. In Cutler, B. L. (Ed.), *Reform of Eyewitness Identification Procedures.* (pp. 87 - 112). Washington, DC: American Psychological Association.

Clark, S. E. & Tunniciliff, J. L. (2001). Selecting lineup foils in eyewitness identification experiments: Experimental control and real-world simulation. *Law and Human Behavior, 25,* 199 – 216. doi:10.1023/A:1010753809988

Clark, S. E., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using receiver operating characteristic analysis. *Current Directions in Psychological Science 23,* 3 – 10. doi:10.1177/0963721413498891

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* (3rd Ed.). New York, New York: Routledge.

Cutler, B. L. & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73,* 281 – 290. doi:10.1037/0021-9010.73.2.281

Davis, J. P., Valentine, T., Memon, A., & Roberts, A. J. (2015). Identification on the street: A field comparison of police street identifications and video line-ups in England. *Psychology, Crime, & Law, 21,* 9 – 27. doi:10.1080/1068316X.2014.915322

Dekle, D. J., Beal, C. R., Elliott, Rogers, & Huneycutt, D. (1996) Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology, 10,* 1 – 12. doi:10.1002/(SICI)1099-0720(199602)10:1<1::AID-ACP354>3.0.CO;2-Y

Dobolyi, D. G. & Dodson, C. S. (2013). Eyewitness confidence in simultaneous and sequential lineups: A criterion shift account for sequential mistaken identification overconfidence. *Journal of Experimental Psychology: Applied, 19,* 345 – 357. doi: 10.1037/a0034596

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

Duncan, M. J. (2006). *A signal detection model of compound decision tasks* (Tech. Rep. Np. No. TR2006-256). Toronto, ON: Defence Research and Development Canada.

Dysart, J. E., Lindsay, R. C. L., & Dupuis, P. R. (2006). Show-ups: The critical issue of clothing bias. *Applied Cognitive Psychology, 20,* 1009 – 1023. doi: 10.1002/acp.1241

Dysart, J. E., Lindsay, R. C. L., Hammond, R., & Dupuis, P. R. (2001). Mug shot exposure prior to lineup identification: Interference, transference, and commitment effects. *Journal of Applied Psychology, 86,* 1280 – 1284. doi:10.1037/0021-9010.86.6.1280

Egan, J. P. (1958). *Recognition memory and the operating characteristic.* (Technical Note AFCRC-TN_58-51, AO-152650). Bloomington: Indiana University Hearing and Communication Laboratory.

Gonzalez, R., Ellsworth, P. C., & Pembroke, M. (1993). Response biases in lineups and showups. *Attitudes and Social Cognition, 64,* 525 – 537. doi: 10.1037/0022-3514.64.4.525

Goodsell, C. A., Neuschatz, J. S., & Gronlund, S. D. (2009). Effects of mugshot commitment on lineup performance in young and older adults. *Applied Cognitive Psychology, 23,* 10.1002/acp.1512

Grant v. *Long Beach* (9th Cir. 2002) 315 F.3d 1081, 1088

Greathouse, S. M. & Kovera, M. B. (2009). Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification. *Law and Human Behavior, 33,* 70 – 82. doi:10.1007/s10979-008-9136-x

Green, D. M. & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* Oxford, England: John Wiley.

Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., & Graham, M. (2012). Showups versus lineups: An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition, 1,* 221 – 228. doi: 10.1016/j.jarmac.2012.09.003

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using receiver operating characteristic analysis. *Current Directions in Psychological Science, 23,* 3 – 10. doi: 10.1177/0963721413498891

Haw, R. M., Dickinson, J. J., & Meissner, C. A. (2007). The phenomenology of carryover effects between show-up and line-up identification. *Memory, 15,* 117 – 127. doi:10.1080/09658210601171672

Horry, R., Palmer, M. A., & Brewer, N. (2012). Backloading in the sequential lineup prevents within-lineup criterion shifts that undermine eyewitness identification performance. *Journal of Experimental Psychology: Applied, 18,* 346 – 360.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237 – 251. doi: 10.1037/h0034747

Lawson, V. Z. & Dysart, J. E. (2014). The showup identification procedure: An exploration of systematic biases. *Legal and Criminological Psychology,* 54 – 68. doi:10.1111/j.2044-8333.2012.02.057.x

Lee, M. D. & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course.* New York, New York: Cambridge University Press.

Levi, A. M. (2012). Much Better than the Sequential Lineup: A 120-Person Lineup. Psychology, Crime & Law, *18,* 631 - 640. doi:10.1080/1068316X.2010.526120

Levi, A. M. (2007). Research note: Evidence for moving to an 84-person photo lineup. *Journal of Experimental Criminology, 3,* 377 – 391. doi:10.1007/s11292-007-042-0

Lindsay, R. C. L., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76,* 741 – 745. doi: 10.1037/0021-9010.76.5.741

Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Leach, A.-M., & Bertrand, M. I. (2009). Sequential lineup presentation: Patterns and policy. *Legal and Criminological Psychology, 14,* 13 – 24. 10.1348/135532508X382708

Lindsay, R. C. L., Martin, R., Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior, 18,* 527 – 541. doi:10.1007/BF01499172

Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21,* 391 – 404. doi:10.1023/A:1024807202926

Lindsay, R. C. L. & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70,* 556 – 564. doi: 10.1037/0021-9010.70.3.556

Loftus, E. F. (1979). *Eyewitness Testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F. & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior, 13,* 585 – 589. doi:10.1016/S0022-5371(74)80011-3

Luus, E. C. A. & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15,* 43 – 57. doi:10.1007/BF01044829

MacMillan, N. A. & Creelman, C. D. (2005). *Detection Theory: A User's Guide. (2nd* Ed.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Malpass, R. S., Tredoux, C. G., Mcquiston-Surrett, D. (2009). Response to Lindsay, Mansour, Beaudry, Leach and Bertrand's 'Sequential lineup presentation: Patterns and policy.' *Legal and Criminological Psychology, 14,* 25 – 30. doi:10.1348/135532508X384094

Manson v. Brathwaite (1977). 432 U.S. 98.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33,* 783 – 792. doi: 10.3758/BF03193074

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18,* 361 – 376. doi:10.1037/a0030609

Mosteller, F. & Tukey, J. W. (1968). Data analysis including statistics. In G. Lindzey & E. Aronson (Eds.), *The Handbook of Social Psychology* (pp. 80 – 203). Reading, MA: Addison Wesley.

Munsterberg, H. (1908/2009). *On the Witness Stand: Essays on Psychology and Crime (Classics in Psychology)*. Greentop Academic Press.

Navon, D. (1992). Selection of lineup foils by similarity to the suspect is likely to misfire. *Law and Human Behavior, 16,* 575 – 593. doi:10.1007/BF01044624

Neil v. Biggers. (1972). 409 U.S. 188.

Palmer, M. A. & Brewer, N. (2012). Sequential lineup presentation promotes less-biased criterion setting but does not improve discriminability. *Law and Human Behavior, 36,* 247 – 255. doi:10.1037/h0093923

Palmer, M. A., Brewer N., & Weber, N. (2012). The information gained form witnesses' responses to an initial "blank" lineup. *Law and Human Behavior, 36*, 439 – 447. doi: 10.1037/h0093939

Pryke, S., Lindsay, R. C. L., Dysart, J. E., & Dupuis, P. (2004). Multiple independent identification decisions: A method of calibrating eyewitness identifications. *Journal of Applied Psychology, 89,* 73 – 84. doi:10.1037/0021-9010.89.1.73

Quinlivan, D. S., Neuschatz, J. S., Cutler, B. L., Wells, G. L., McClung, J., & Harker, D. L. (2012). Do pre-admonition suggestions moderate the effect of unbiased lineup instructions? *Legal and Criminological Psychology, 17,* 165 – 176. doi:10.1348/135532510X533554

Rosenthal, R. & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences, 1,* 377 – 415. doi:10.1017/S0140525X00075506

R. v. Mann. (2004). 3 S.C.R. 59.

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

R. v. Watson, (1996) O.J. No. 2695

Smith, A. M., Bertrand, M. I., Lindsay, R. C. L., Kalmet, N., Grossman, D., & Provenzano, D. (2014). The impact of multiple show-ups on eyewitness decision-making and innocence risk. *Journal of Experimental Psychology: Applied.* Advance online publication. http: dx.doi.org/10.1037/xap0000018

Smith, A. M. & Dufraimont, L. (2014). Safeguards against wrongful conviction in eyewitness identification cases: Insights from empirical research. *Canadian Criminal Law Review, 18,* 199 – 219.

Smith, A. M., Leach, A.-M., Cutler, B. L. (2013) Facilitating accuracy in showup identification procedures: The effect of the presence of stolen property on identification accuracy. *Applied Cognitive Psychology, 27,* 216 – 221. doi: 10.1002/acp.2898

Smith, S. M. (1988). Environmental context – dependent memory. In Davies, G. M. & Thomson, D. M. (Ed.) *Memory in Context: Context in Memory* (pp. 231 – 244). New York, NY: John Wiley & Sons.

Smith, S. M. (1994). Theoretical principles of context-dependent memory. In Morris, P., and Gruneberg, M. (Eds.). *Aspects of Memory* (*2nd edition*): *Theoretical Aspects* (pp.168 – 195). Routledge Press.

Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31,* 137 – 149. doi: 10.3758/BF03207704

Steblay, N. K. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21,* 283 – 297. doi:10.1023/A:1024890732059

Steblay, N. K. (2013). Lineup instructions. In Cutler, B. L. (Ed.), *Reform of Eyewitness Identification Procedures.* (pp. 65 - 86). Washington, DC: American Psychological Association.

Steblay, N. K., Dysart, J. E., 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,* 459 – 473. doi:10.1023/A:1012888715007

Steblay, N. K., Dysart, J. E., 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,* 523 – 540. doi:10.1023/A:1025438223608

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law, 17,* 99 – 139. doi:10.1037/a0021650

Stovall v. Denno (1967). 388 U.S. 293.

Technical Working Group for Eyewitness Evidence. (1999). Eyewitness evidence: A guide for law enforcement. Washington, DC: United States Department of Justice. Retrieved November 15, 2009, from: https://www.ncjrs.gov/pdffiles1/nij/178240.pdf

The National Registry of Exonerations. (2015). https://www.law.umich.edu/special/exoneration/Pages/about.aspx

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods, 6,* 371 – 386. doi: 10.1037/1082-989X.6.4.371

Tunnicliff, J. L. & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions? *Law and Human Behavior, 24,* 231 – 258. doi:10.1023/A:1005463020252

Valentine, T., Davis, J. P., Memon, A., & Roberts, A. (2012). Live showups and their influence on a subsequent video line-up. *Applied Cognitive Psychology, 26,* 1 – 23. doi:10.1002/acp.1796

Wageenar, W. A. & Veefkind, N. (1992). Comparison of one-person and many-person lineups: A warning against unsafe practices. In Lösel, F., Bender, D., & Bliesener, T., (Eds.). *Psychology and Law: International Perspectives.* (pp. 275 - 285). Oxford, England: Walter De Gruyter.

Washington v. Lambert. (1996). 98 F. 3d 1181.

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36,* 1546 – 1557. doi: 10.1037/0022-3514.36.12.1546

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14,* 89 – 103. doi: 10.1111/j.1559-1816.1984.tb02223.x

Wells, G. L. (2014). Eyewitness identification: Probative value, criterion shifts, and policy regarding the sequential lineup. *Current Directions in Psychological Science, 23,* 11 – 16. doi:10.1177/0963721413504781

Wells, G. L., Leippe, M. R., & Ostrom, T. M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior, 2,* 285 – 293. doi: 10.1007/BF01039807

Wells, G. L. & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88,* 776 – 784. doi: 10.1037/0033-2909.88.3.776

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,* 106 – 117. doi: 10.1177/0146167290161008

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78,* 835 – 844. doi:10.1037/0021-9010.78.5.835

Wells, G. L., Smalarz, L., & Smith, A. M. (In Preparation). ROC analysis of lineups does not measure underlying discriminability and has limited value. *Journal of Applied Research in Memory and Cognition.*

Wells, G. L., Steblay, N. K., & Dysart, J. E. (2012). Eyewitness identification reforms: Are suggestiveness-induced hits and guesses true hits? *Perspectives on Psychological Science, 7,* 264 – 271. doi:10.1177/1745691612443368

Wells, G. L. & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99,* 320 – 329. doi: 10.1037/0033-2909.99.3.320

Wells, G. L. & Wilford, M. M. (2013). Eyewitness system variables. In Cutler, B. L. (Ed.) *Reform of Eyewitness Identification Procedures* (pp. 23 – 43). Washington, DC: American Psychological Association.

Wells, G. L. & Windschitl, P. D. (1999). Stimulus Sampling and Social Psychological Experimentation. *Personality and Social Psychology Bulletin*, *25*(9), 1115-1125. doi: 10.177/01461672992512005

Wells, G. L., Yang, Y., & Smalarz, L. (2015). Eyewitness identification: Bayesian information gain, base-rate effect equivalency curves, and reasonable suspicion. *Law and Human Behavior, 39,* 99 – 122. doi:10.1037/lhb0000125

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Wooten, A., Goodsell, C. A., Carlson, C. A. (2015). Effect of retention interval on showup and lineup performance. *Journal of Applied Research in Memory and Cognition, 4,* 8 – 14. doi:10.1016/j.jarmac.2017.07.003

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Key, K. N., & Goodsell, C. A. (2015). Do the clothes make the criminal? The influence of clothing match on identification accuracy in showups. *Journal of Applied Research in Memory and Cognition, 4,* 36 – 42. doi:10.1016/j.jarmac.2014.12.003

Whipple, G. M. (1909). The observer as reporter: A survey of the psychology of testimony. *Psychological Bulletin, 6,* 153 – 170. doi:10.1037/h0071084

Wigmore, J. H. (1909). Professor Muensterberg and the psychology of testimony: Being a report of the case of Cokestone v. Muensterberg. *Illinois Law Review, 3,* 399 – 434.

Wixted, J. T., Gronlund, S. D., & Mickes, L. (2014). Policy regarding the sequential lineup is not informed by probative value but is informed by receiver operating characteristic analysis. *Current Directions in Psychological Science, 23,* 17 – 18. doi:10.1177/0963721413510934

Wixted, J. T. & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives on Psychological Science, 7,* 275 – 278. doi:10.1177/1745691612442906

Yarmey, A. D., Yarmey, A. L., & Yarmey, M. J. (1994). Face and voice identifications in showups and lineups. *Applied Cognitive Psychology, 8,* 453 – 464. doi: 10.1002/acp.2350080504

Yarmey, A. D., Yarmey, M. J., & Yarmey, A. L. (1996). Accuracy of eyewitness identifications in showups and lineups. *Law and Human Behaviour, 20,* 459 – 477. doi: 10.1007/BF01498981