

The Houston Police Department Eyewitness Identification Experiment:

# ANALYSIS AND RESULTS

A Report to the Houston Police Department





The Houston Police Department Eyewitness Identification Experiment:  
**Analysis and Results**

A Report to the Houston Police Department

William Wells  
College of Criminal Justice  
Sam Houston State University

July 21, 2014

Copyright 2014

## ACKNOWLEDGMENTS

Conducting high-quality field experiments in the criminal justice system presents significant obstacles. Criminal justice scholars have learned it is essentially impossible to conduct good field experiments without the support and cooperation of the people who work in the system every day. Many individuals in the Houston Police Department (HPD) supported this experiment and helped design, launch, and complete this study. This space is not adequate to acknowledge the dedication and support these people demonstrated. The following individuals played an instrumental role in launching and completing the experiment:

Martha Montalvo, Executive Assistant Chief

George Buenik, Assistant Chief

Donald McKinney, Assistant Chief

Lori Bender, Captain

Mark Holloway, Captain

James Jones, Captain

Heather Morris, Lieutenant

Stephen Morrison, Sergeant

Dennis Schoppe, Sergeant

The men and women of the Robbery Division are to be commended for embracing the demands of a unique experiment that will benefit HPD and police departments across the country.

Three graduate students in the Department of Criminal Justice and Criminology served as research assistants on the project. Brad Campbell, Yudu Li, and Stryker Swindle worked many long hours on various aspects of the project and they were instrumental to its success.

Four distinguished scholars served on the project advisory board and provided valuable feedback about the design of the study, the treatment conditions that were tested, the data collection procedures, and project reports.<sup>1</sup> The advisory board has collective expertise in the area of research on eyewitness identification procedures and on the nuances of conducting successful field studies with police agencies. The advisory board included:

Steven Clark, Ph.D., University of California at Riverside

Jennifer Dysart, Ph.D., John Jay College of Criminal Justice

Edward Maguire, Ph.D., American University

Jeremy Wilson, Ph.D., Michigan State University

---

<sup>1</sup> Vincent Webb, Ph.D., of the Department of Criminal Justice and Criminology at Sam Houston State University, also provided valuable feedback and advice throughout the course of the study.

# CONTENTS

|   |           |
|---|-----------|
| <b>Executive Summary</b> .....                                | <b>1</b>  |
| <b>Introduction</b> .....                                     | <b>3</b>  |
| <b>Experimental Conditions and Procedures</b> .....           | <b>3</b>  |
| Randomization Procedures.....                                 | 4         |
| Data Collection and Data Entry Procedures .....               | 6         |
| Fidelity .....  | 6         |
| <b>Results</b> .....  | <b>7</b>  |
| Data Overview .....   | 7         |
| Sub-Sample Analyses.....                                      | 9         |
| Analysis #1 Patterns of results among 718 photo spreads ..... | 10        |
| Analysis #1 Results.....                                      | 11        |
| Analysis #1 Summary .....                                     | 15        |
| Analysis #2 Patterns of results among 830 photo spreads ..... | 15        |
| Analysis #2 Results.....                                      | 16        |
| Analysis #2 Summary .....                                     | 19        |
| <b>Conclusion</b> .....                                       | <b>20</b> |
| <b>References</b> .....                                       | <b>21</b> |
| <b>Appendix A</b> .....                                       | <b>23</b> |
| <b>Appendix B</b> .....                                       | <b>26</b> |
| Lineup Fairness Assessment.....                               | 26        |
| Methods .....   | 26        |
| Findings .....  | 27        |
| <b>Appendix C</b> .....                                       | <b>30</b> |
| Analysis #1 Multivariate models .....                         | 30        |
| Analysis #2 Multivariate models .....                         | 34        |

## TABLES AND FIGURES

|           |   |    |
|-----------|---|----|
| Table 1.  | Procedures used during the experiment. ....   | 8  |
| Table 2.  | Identification results in the full sample of photo spreads and lineups.....                     | 9  |
| Table 3.  | Differences and similarities between treatment groups. ....                                     | 11 |
| Table 4.  | Differences in selecting a photo across experimental treatments.....                            | 12 |
| Table 5.  | Differences in selecting a photo, a filler, and the suspect across experimental treatments..... | 13 |
| Figure 1. | Suspect and filler selections among cases in which a selection was made.....                    | 13 |

|            |   |    |
|------------|---|----|
| Table 6.   | Differences in selecting a photo, a filler, and the suspect across experimental treatments: Disaggregating positive from tentative identifications.....   | 14 |
| Table 7.   | Differences and similarities between treatment groups: Sample of 830 cases....  | 16 |
| Table 8.   | Differences in selecting a photo across experimental treatments.....  | 17 |
| Table 9.   | Differences in selecting a photo, a filler, and the suspect across experimental treatments.....   | 17 |
| Figure 2.  | Suspect and filler selections among cases in which a selection was made: Sample of 830 cases.....   | 18 |
| Table 10.  | Differences in selecting a photo, a filler, and the suspect across experimental treatments: Disaggregating positive from tentative identifications.....   | 19 |
| Table A-1. | Identification procedures and experimental protocol violations.....   | 24 |
| Table A-2. | Identification procedures assigned and actually used.....   | 24 |
| Table A-3. | Experimental treatment that should have been used and treatments actually used.....   | 25 |
| Table B-1. | Blinded simultaneous photo spreads: Percentage of mock witnesses who identified the suspect.....  | 27 |
| Table B-2. | Blinded sequential photo spreads: Proportion of mock witnesses who identified the suspect.....  | 28 |
| Table B-3. | Blind simultaneous photo spreads: Proportion of mock witnesses who identified the suspect.....  | 28 |
| Table B-4. | Blind sequential photo spreads: Proportion of mock witnesses who identified the suspect.....  | 28 |
| Table B-5. | Average bias and effective size of the sampled photo spreads.....   | 28 |
| Table B-6. | One-Way ANOVA for bias and size across of experimental conditions.....  | 29 |
| Table C-1. | Multivariate logistic regression results: Effects of photo spread administration method on the likelihood a photo was selected.....                       | 30 |
| Table C-2. | Multivariate multinomial regression results: Effects of photo spread administration method on selection results.....                                      | 31 |
| Table C-3. | Multivariate multinomial regression results: Effects of photo spread administration method on selection results, disaggregated by witness confidence..... | 33 |
| Table C-4. | Multivariate logistic regression results: Effects of photo spread administration method on the likelihood a photo was selected.....                       | 34 |
| Table C-5. | Multivariate multinomial regression results: Effects of photo spread administration method on selection results.....                                      | 35 |
| Table C-6. | Multivariate multinomial regression results: Effects of photo spread administration method on selection results, disaggregated by witness confidence..... | 37 |

## EXECUTIVE SUMMARY

**Purpose:** The purpose of this experiment was to compare four different methods of administering photo spreads and two methods of administering lineups to eyewitnesses during robbery investigations: blinded simultaneous, blinded sequential, blind simultaneous, and blind sequential. This is the second report as part of this project. The first report (Wells, 2013) provides detailed information about the research design and data collection instruments.

**Method:** A randomized experimental design was used to assign identification procedures to criminal investigations. At the end of each identification procedure investigators and witnesses completed a survey to capture information about the robbery and the identification procedure.

**Data:** Investigators returned surveys from 1,096 identification procedures conducted between January 22 and December 5, 2013.

**Experimental Fidelity:** Overall, investigators adhered to experimental protocols in a large portion of photo spread procedures. Deviations from the experimental protocol occurred more often when a sequential method was assigned to be used. Deviations from experimental protocols were driven, in large part, by photo spread procedures conducted as part of serial robbery investigations.

**Equivalent Experimental Groups:** The four experimental groups differed in some important ways. For this reason it was necessary to estimate statistical models that controlled for differences between the four treatment groups.

### Results:

Overall, there were few meaningful differences between the methods in terms of filler selection rates (i.e., known selection errors) and suspect selection rates. These results consistently emerged from different forms of analyses. The patterns of results did not indicate one method or set of methods consistently performs better than others in terms of reducing filler selections.

Blinded simultaneous photo spreads reduced the chances that any photo was selected compared to the other three procedures.

The highest suspect selection rate was produced with the blinded sequential procedure.

Overall, suspects were positively identified in approximately 20% of photo spreads while fillers were positively identified in less than 5% of photo spreads.

The most common result of a photo spread was that the witness did not make a selection.

Future analyses will explore additional policy-relevant relationships.



## INTRODUCTION

The first report from the Houston Police Department (HPD) Eyewitness Identification Experiment provides a detailed description of the methods used in this experiment (Wells, 2013). The details of the research design will not be reiterated in this report, but a brief overview is provided. The purpose of this, the second report, is to describe the data and results of the experiment. The data generated through this experiment are rich and additional analyses will be on-going; the analyses presented here represent the first steps. As findings continue to emerge they will be provided to HPD for review.

Data collection lasted from January 22 through December 5, 2013. A randomized experimental design was used because this design is well-suited for estimating the effects of a treatment, or intervention, on the outcome of interest (Shadish, Cook, & Campbell, 2002). A randomized design is strong because it controls for relevant factors that may influence the outcome and isolates the effect of the treatment. In this case the treatment represents different methods of showing photo spreads, live lineups, and video lineups to witnesses.<sup>2</sup> Despite the strength of randomized experimental designs, they are challenging to implement in field settings with criminal justice system agencies (see, for instance, Feder, Jolin, & Feyerherm, 2000; Sherman, Schmidt, & Rogan, 1992; Weisburd, 2000; Weisburd, Petrosino, & Mason, 1993; Wells, Steblay, & Dysart, 2011). The experiment was designed with care to increase the chances that experimental protocols would be followed and so that it was possible to measure adherence to protocols. Assessing adherence to experimental protocols is common in field experiments with criminal justice system agencies (see, for instance, McGarrell, Chermak, Weiss, & Wilson, 2001; Mazerolle, Roehl, & Kadleck, 1998; Sherman, et al., 1992; Wells, et al., 2011). The next sections provide a brief overview of the study design. Readers are encouraged to consult the first project report that provides additional details (Wells, 2013).

## EXPERIMENTAL CONDITIONS AND PROCEDURES

The experiment tested four different methods of showing photo spreads:

- blind sequential
- blind simultaneous
- blinded sequential
- blinded simultaneous

The experiment also tested two methods of showing individuals in live and video lineups.<sup>3</sup>

- blind sequential
- blind simultaneous

The Texas Code of Criminal Procedure 38.20 requires law enforcement agencies to use procedures that minimize the influence that an administrator can have on witnesses when they view photo spreads and individuals in lineups. The methods tested in this experiment are consistent with the

<sup>2</sup> The terms “witness” and “eyewitness” are used here in reference to individuals who view photo spreads and lineups. Witnesses include individuals who are bystanders and who are witnesses because they are also victims.

<sup>3</sup> Video lineups are video-recorded live lineups. All live lineups are video recorded so the procedures are captured and can be reviewed. Video lineups can be shown to witnesses if they are unable to attend the live lineup. Throughout this report the generic term “lineup” is used to refer to both live and video lineups.

requirements of the Texas Code of Criminal Procedure and represent methods likely to be used by law enforcement agencies across the state. Each of these procedures is consistent with HPD eyewitness identification policy. Even though these methods are permissible under HPD policy, investigators may not have been familiar with the nuances of each technique and what was required as part of the experiment. All HPD robbery investigators were trained on the experimental protocols and the use of the different identification procedures in late 2012. Investigators were provided with a user-friendly instruction booklet they could reference throughout the study period. This was intended to provide investigators with information about the procedures in the event that questions arose when they were preparing to conduct an identification procedure.

With a simultaneous presentation method the eyewitness views all photos or lineup members at the same time. In a sequential presentation method, photos or lineup members are viewed one at a time. With a blind procedure the primary investigator who knows the identity of the suspect prepares the photo spread but an investigator with no knowledge of the suspect's identity conducts the viewing procedure with the witness. Finally, the primary investigator conducts the viewing in a blinded procedure, but a mechanism is used to prevent the investigator from 1) knowing the suspect's position in the photo spread or lineup and 2) knowing which photo or individual the witness is viewing. The detailed procedures HPD robbery investigators used for this experiment are described in Wells (2013).

Administering photo spreads and lineups in a blind manner is relatively straightforward. An independent administrator who does not know the identity of the suspect and who does not know the position of the suspect in the photo spread or lineup administers the procedure. Blinded procedures can be accomplished in more than one way but the primary intent is to reduce the chances for the police administrator to influence the witness's selection. One method that is widely available to police agencies relies on manila folders or envelopes. The experiment tested a folder blinding method with sequential and simultaneous techniques. These techniques represent methods likely to be used by many police agencies, including HPD. The experiment is strong in this respect because the techniques being tested are currently used or being considered for use by HPD and other police agencies in Texas. The first project report provides a complete description of the blind and blinded procedures used in the experiment (Wells, 2013). The admonishment form HPD used as of January 2013 was used throughout the experiment (see Wells, 2013).

## Randomization Procedures

"Work," "office," and "pending" reports received by the Robbery Division were randomly assigned to receive one of the four experimental procedures: blind sequential, blinded sequential, blind simultaneous, or blinded simultaneous.<sup>4</sup> "Work," "office," and "pending" cases were assigned to a treatment condition because these are the case classifications that generate investigative activities, including photo viewings and lineups. In cases of live and video lineups a blind administrator administers to the lineup; the investigator working the case is not present when the eyewitness is viewing the live lineup. The characteristics of the case determined whether a live lineup or a video lineup would be conducted.<sup>5</sup>

When a robbery report is routed to the Robbery Division it is automatically assigned an HPD case number (referred to as an "incident number" in HPD databases) by the computerized information management system. For purposes of the study, the case number determined which experimental

<sup>4</sup> Please see the first project report for a complete description of the robbery case typology (Wells, 2013).

<sup>5</sup> When case managers classify and assign cases to investigators they do not know if it will be possible to conduct a video or live lineup during the investigation. Thus, it is not possible to randomly assign "photo viewing" and "lineup" treatments to the criminal cases.

procedure was to be used with all eyewitness identification procedures in the investigation. The study randomly assigned *investigations* to a treatment condition rather than identification procedures. The last two digits of the case number represent the year the event occurred (i.e., 12 for cases occurring in 2012) and did not determine which procedure was assigned to the investigation. The third digit and the fourth digit from the last digit represent sequential numbers that are automatically generated by HPD data systems. The third from the last digit determined whether a simultaneous or a sequential method was assigned to the investigation. When the third from the last digit was an even number then a simultaneous procedure was to be used and when the digit was an odd number a sequential procedure was to be used. The fourth from the last digit determined whether a blinded or a blind method would be used during the investigation. When the fourth from the last digit was an even number then robbery investigators were to use a blinded procedure and when the digit was an odd number the investigators were to use a blind procedure. Robbery investigators do not have discretion in determining the case number that is assigned to a case and they do not control which cases are assigned to them by case managers. Thus, random assignment of investigations to treatment conditions should be expected to create groups of identification procedures and criminal cases that are roughly equivalent in important ways.<sup>6</sup>

All photo spreads and video and live lineups conducted in a given case were to utilize the same method that is assigned to the case number. For instance, when a case number dictated that a simultaneous method should be used and the detective showed five photo spreads, all five photo spreads were to be shown with the simultaneous method. One purpose of assigning investigations to treatment conditions was to ensure that all identification procedures conducted as part of an investigation were consistent. This was intended to minimize potential confusion with jurors, attorneys, and judges if a case proceeded to trial. An exception was when an investigation involved live or video lineups and the experimental condition was assigned to be “blinded.” In this case it was not possible to conduct a “blinded” lineup (i.e., the investigator knows the identity of the suspect but does not know the suspect’s position in the lineup) so all lineups were conducted blind (i.e., the investigator does not know the identity of the suspect and does not know the suspect’s position in the lineup). In this hypothetical investigation the photo spreads would be conducted blinded and lineups would be conducted blind. The exception to this rule was when the first identification procedure conducted as part of an investigation is a live or a video lineup. In this situation the subsequent identification procedures were to be conducted blind to create consistency of procedure within the investigation.

Serial events introduced complications for the randomization procedure. Serial robberies are common events and entail the connection of multiple, seemingly distinct robbery incidents. When case managers and investigators draw connections between three or more robbery events then these are referred to as a “serial investigation.” Serial robberies entail three or more robberies carried out by the same individual or group of individuals. Two robberies connected together are referred to as “related cases” rather than “serial cases.” Most important for the experiment, HPD robbery investigations treat these serials and related cases as single investigations. Consistent with the procedures described above (i.e., all procedures conducted as part of a single investigation should be conducted in the same way), all eyewitness identification procedures conducted as part of a serial investigation and a related-case investigation were to be conducted the same way. If no procedure had been conducted with any witnesses when the serial was identified, the treatment procedure assigned to the case in the serial that occurred first determined the presentation method used. In these situations, all eyewitness procedures in the cases that are part of a serial were to rely on the

---

<sup>6</sup> See Shadish, et al. (2002) for an explanation about why random assignment is uniquely valuable for understanding causal relationships between treatments and outcomes. The analysis described below compared the four groups of cases to determine if randomization procedures generated roughly equivalent groups.

same administration procedure (i.e., simultaneous or sequential). When serial cases were connected together and different viewing procedures had been used in these cases, a lieutenant or captain decided which case in the serial is most central to the investigation; all future viewing procedures were to utilize the method that was used in that most central case. The decision to assign a single showing method to all cases in the serial was made to avoid potential confusion and concern that the use of different methods may cause for jurors, attorneys, and judges should a serial or a related-case investigation advance to court proceedings.<sup>7</sup>

When case managers assigned cases to investigators they used color-coded file folders that indicated which procedure was to be used during the investigation. In addition, case managers stamped the procedure to be used on the front page of the file folder. Red folders were used for blinded simultaneous procedures, green folders for blind simultaneous procedures, blue folders for blinded sequential procedures, and yellow folders were used for blind sequential procedures.

### Data Collection and Data Entry Procedures

Investigators and witnesses completed surveys at the conclusion of each photo spread and lineup procedure. These surveys provided information about the witness's view of the offender(s) during the crime and about the identification procedure (see Wells, 2013 for additional details). Investigators reported information on the survey that was available to them at the time of the identification procedure. At the conclusion of the identification procedures both surveys were placed in a large envelope and returned to the Robbery Division at headquarters. Members of the SHSU research team drove to HPD headquarters and picked up completed surveys throughout the course of the experiment.

Survey data were entered into an electronic database (SPSS v20.0) at SHSU. All investigator and witness survey data were entered into the database two times by different members of the research team. Once data for each witness and investigator survey were entered twice, the information was compared for consistency. When inconsistencies were identified a member of the research team examined the original survey and corrected information in the SPSS database. This procedure ensured that data entry errors were identified and corrected.

### Fidelity

An important challenge with field experiments in criminal justice is ensuring that research protocols are followed. Field experiments impose new protocols, disrupt work routines, and sometimes remove discretion from certain decisions. For these reasons it is important to monitor study fidelity. In this experiment it is important that 1) cases were randomly assigned to administration procedures; 2) investigators used the administration procedure that was assigned; and 3) detective and witness surveys were completed accurately. It is critical that the four groups of cases (i.e., cases in each of the four treatment conditions) are, in the aggregate, equivalent (see Shadish et al., 2002). The analysis described below examined the equivalence of the four groups of cases in terms of relevant variables contained in the investigator survey.

---

<sup>7</sup> A photo spread procedure conducted as part of a serial investigation was considered a violation of the experimental protocol if it deviated from the method that would be assigned to the earliest occurring robbery in the serial. If the procedure was consistent with what would have been assigned to the first robbery that occurred in the serial, then this was not considered a protocol violation.

---

## RESULTS

### Data Overview

Investigators returned surveys that captured information about 1,096 identification procedures that occurred between January 22 and December 5, 2013. It is not feasible to extract from HPD data systems the total number of identification procedures conducted during this time period. Thus, it is not possible to determine the percentage of procedures represented by the survey data. In the analyses described below some of these surveys were excluded for a variety of reasons. Twenty-seven surveys were eliminated because the lead investigator in the case did not sign an informed consent form required by Sam Houston State University to participate in the experiment. Twelve cases were eliminated because the surveys were missing key pieces of information, including the incident number, the method used during the procedure, and the selection result. When these 39 cases were eliminated, this produced a sample of 1,057 identification procedures that are described below.

The sample of 1,057 identification procedures entailed 29 live lineups, 31 video lineups, and 997 photo spreads. Approximately 6% of the procedures conducted during this time period were lineups and 94% were photo spreads. The set of 1,057 identification procedures were conducted as part of 390 separate investigations: 330 single-incident investigations and 60 serial investigations and investigations with a related case. Serial investigations involve three or more incidents and related-case investigations involve two incidents. Single-incident investigations entailed, on average, approximately two identification procedures while serial and related-case investigations entailed, on average, slightly more than six identification procedures.

To reiterate, the experiment tested four procedures for administering photo spreads:

- *Blinded simultaneous*: photos are shown at the same time on a single sheet of paper by an investigator who knows the identity of the suspect but does not know the suspect's position in the photo spread.
- *Blinded sequential*: photos are shown individually (one at a time) by an investigator who knows the identity of the suspect but does not know the suspect's position in the photo spread.
- *Blind simultaneous*: photos are shown at the same time on a single sheet of paper by an investigator who does not know the identity of the suspect and does not know the suspect's position in the photo spread.
- *Blind sequential*: photos are shown individually (one at a time) by an investigator who does not know the identity of the suspect and does not know the suspect's position in the photo spread.

Table 1 shows the procedures used during the experiment. Two lineup procedures conducted in a blinded manner represent protocol violations because live and video lineups must be conducted in a blind manner (i.e., the investigator who administers the lineup must not know the identity of the suspect and must not know the location of the suspect in the lineup).

Table 1. Procedures used during the experiment.

|                      | <b>Photo Spreads</b> | <b>Live Lineups</b> | <b>Video Lineups</b> | <b>Total</b>       |
|----------------------|----------------------|---------------------|----------------------|--------------------|
| Blinded Simultaneous | 289 (29%)            | 0                   | 1 (3%)               | 290 (27%)          |
| Blinded Sequential   | 221 (22%)            | 1 (3%)              | 0                    | 222 (21%)          |
| Blind Simultaneous   | 270 (27%)            | 18 (62%)            | 17 (55%)             | 305 (29%)          |
| Blind Sequential     | 217 (22%)            | 10 (35%)            | 13 (42%)             | 240 (23%)          |
| <b>Total</b>         | <b>997 (100%)</b>    | <b>29 (100%)</b>    | <b>31 (100%)</b>     | <b>1057 (100%)</b> |

When investigators completed a survey as part of the experiment they were asked to indicate which individual(s) in the lineup or photo spread the witness identified during the procedure. Specifically, they were asked to report the position (number 1 through 6) the witness selected (see Wells, 2013 Appendix A). Investigators were also asked, later in the survey, to record the final selection result from the lineup or photo spread (i.e., no person was selected, the suspect was identified, or a filler was identified).

On its face, measuring the result of an identification procedure is a seemingly straightforward task. Witnesses can identify a suspect, filler, or not select anyone from the lineup or photo spread. In practice this is complicated. Not surprisingly, the survey data show that in some cases witnesses found it difficult to make a selection (see also Wells et al., in press). In 92 procedures (9%), the survey data indicate that witnesses, at least tentatively, indicated that multiple members of the photo spread or lineup could be the perpetrator. In instances when multiple individuals were selected we made the assumption that the investigator’s summary judgment about the final selection outcome measured the final result. This assumption is grounded in the idea that an investigator is present with the witness during the procedure and is able to best determine the final result. In some cases this was straightforward because the witness selected two or three fillers during the procedure and this was coded a filler selection. When witnesses indicated a suspect and a filler were identified, members of the research team, in collaboration with HPD personnel, consulted information in the HPD On-Line Offense (OLO) system and consulted other items in the survey to make a determination about the final selection result. A small number of cases (n = 18, see Table 2) were coded “other” because it was not possible to determine the final selection result based on available information.

Table 2 shows the results of *all* identification procedures in the experiment. Suspect identifications occurred most frequently (n = 415) followed by no selections (n = 369) and filler identifications (n = 255). When robbery investigators read the admonishment form to a witness at the beginning of an identification procedure they explain that HPD uses three levels of certainty the witness can use if they identify someone. The highest degree of certainty is a “positive” identification, the mid-level of certainty is a “strong tentative” selection, and “weak tentative” selections are used to classify instances with the lowest degree of certainty. Table 2 shows positive suspect identifications occurred in 26% of cases while positive filler identifications occurred infrequently, in 4% of cases. By contrast “tentative” identifications were more common when a witness believed a filler was the perpetrator. When some selection was made, a positive suspect selection occurred in 41% of cases and a positive filler selection occurred in 6% of cases.

Table 2. Identification results in the full sample of photo spreads and lineups.

|                           | Count       | Percentage | Valid Percentage <sup>a</sup> |
|---------------------------|-------------|------------|-------------------------------|
| <i>No Selection Made</i>  | 369         | 35         |                               |
| <i>Suspect Identified</i> | 415         |            |                               |
| Positive                  | 279         | 26         | 41                            |
| Strong Tentative          | 95          | 9          | 14                            |
| Weak Tentative            | 41          | 4          | 6                             |
| <i>Filler Identified</i>  | 255         |            |                               |
| Positive                  | 38          | 4          | 6                             |
| Strong Tentative          | 70          | 7          | 10                            |
| Weak Tentative            | 147         | 14         | 21                            |
| <i>Other</i>              | 18          | 2          | 3                             |
| <b>Total</b>              | <b>1057</b> | <b>100</b> | <b>100</b>                    |

<sup>a</sup> Percentages in this column are based on 688 cases in which a selection was made.

### Sub-Sample Analyses

The results presented above provide an overview of *all cases* combined together, but to understand whether there are different selection results across the different administration procedures it is important to examine sub-samples of cases. To produce the sample of cases used in the primary analyses, several types of cases are excluded. First, the primary analyses focus on photo spreads (lineups are excluded) because these are the most common procedures used by HPD and because lineups do not allow for a comparison of blind and blinded procedures (all lineups must be conducted in a blind manner). Second, photo spreads involving witnesses and suspects who are non-strangers are excluded. The idea is that photo spread procedures should make very little difference when the witness is familiar with the suspect. Third, instances in which a suspect was shown to the same witness multiple times are excluded. These cases are excluded because witnesses may be more likely to select an individual from a photo spread during subsequent viewings simply because they were exposed to a suspect in a previous identification procedure. Fourth, cases are excluded if the final selection result from the procedure is not available in the data or is unclear. Finally, cases are excluded when an identification procedure deviated from the procedure that was supposed to be used under the random assignment protocol. For instance, if a case was assigned to use the blinded simultaneous procedure but a different procedure was used then this case is excluded from the analysis. These cases are referred to as “experimental protocol violations.” The analyses reported below examined cases that did and did not involve experimental protocol violations in order to understand how sensitive the findings are to the inclusion and exclusion of these cases.

Appendix A provides detailed information about the sub-sample of cases that are used in the analyses presented below. It is important to note that experimental protocol violations do *not* appear to be evenly distributed across the experimental groups (see Table A-1). For this reason it is important to compare the four experimental groups in terms of relevant variables to understand the degree to which they are similar and different. These results are presented below. All analyses described below were carried out using SPSS v20.0.

## Analysis #1 Patterns of results among 718 photo spreads

The first set of results is based on the analysis of 718 photo spreads (see Appendix A). This sub-sample only includes robberies involving strangers, witnesses who had not previously viewed a photo spread with the suspect, and photo spreads that followed the experimental protocol that should have been used during the investigation.

The first step in the analysis determined the degree to which the randomization procedure generated four treatment groups that are roughly equivalent in terms of relevant variables (see Wells, 2013 for information about the measurement of these indicators). Table 3 shows the results of these comparisons. The experimental groups differ in some ways. Interpreters were used slightly more often in blind sequential cases (i.e., photos are shown one at a time by an investigator who does not know the identity of the suspect and does not know the suspect's position in the photo spread). Corroborating evidence was available more often in blinded sequential cases (i.e., photos are shown one at a time by an investigator who knows the identity of the suspect but does not know the suspect's position in the photo spread). Witnesses were reported to be under the influence of drugs or alcohol at the time of the robbery less often in blind simultaneous cases. Witnesses reported to have seen a photo of the perpetrator between the crime and the time of the identification procedure less often in blind sequential cases. Among witnesses who wore glasses, it was more often the case that they were not wearing them during the crime in blind sequential cases. Finally, blind simultaneous procedures (i.e., photos are shown all at the same time on a single sheet of paper by an investigator who does not know the identity of the suspect and does not know the suspect's position in the photo spread) were conducted more often in police facilities and less often in public settings and the witness's workplace or school.

These differences between the experimental groups are controlled in multivariate models that compare the four groups in terms of witness identification results. It is necessary to control for these differences in statistical models because the differences may affect the patterns of results and lead to inaccurate conclusions about similarities or differences between the photo spread methods.

The four groups were also compared in terms of photo spread fairness (or bias). Routine practice is for the lead investigator in a case to create photo spreads by selecting the suspect's photo and five filler photos. Investigators enter descriptive information into the computerized photo system and a set of photos containing those descriptors is generated. The investigator then selects the filler photos to include in the photo spread. This procedure was followed during the course of the experiment in order to replicate normal operating procedures that exist outside of the context of the experiment. Investigators know which identification method they are supposed to use prior to creating photo spreads. This procedure presents a potential threat to the validity of the study because investigators may create photo spreads that are different across the four experimental groups. In other words, investigators may intentionally or unintentionally create photo spreads that make it more or less difficult for the witness to select a suspect in some of the experimental groups but not others.<sup>8</sup> If the photo spreads or lineups within a specific treatment condition are more or less challenging than those in the other treatment conditions, this makes it difficult to draw conclusions about whether the treatments are responsible for any observed differences between the groups in terms of witness selection results. Appendix B describes the procedures that were used to compare the quality of photo spreads across the four treatment groups. The results obtained from a sample of 60 photo spread procedures suggest that the treatment groups do not differ in terms of the quality of photo spreads.

---

<sup>8</sup> The author thanks a member of the project advisory board for stressing the importance of this threat to experimental fidelity.

Table 3. Differences and similarities between treatment groups.

| Case Characteristics                | Photo Spread Administration Method |                               |                               |                             | Statistical Test       |
|-------------------------------------|------------------------------------|-------------------------------|-------------------------------|-----------------------------|------------------------|
|                                     | Blinded Simultaneous<br>N = 194    | Blinded Sequential<br>N = 175 | Blind Simultaneous<br>N = 187 | Blind Sequential<br>N = 162 |                        |
| Interpreter Used                    | 6 (3%)                             | 13 (7%)                       | 16 (9%)                       | 23 (14%)                    | $\chi^2=14.82 p=.002$  |
| Corroborating Evidence              | 121 (62%)                          | 160 (91%)                     | 130 (70%)                     | 106 (65%)                   | $\chi^2=45.73 p=.000$  |
| Witness Under Influence             | 17 (9%)                            | 13 (7%)                       | 3 (2%)                        | 13 (8%)                     | $\chi^2=10.00 p=.019$  |
| Witness Saw Photo                   | 21 (11%)                           | 16 (9%)                       | 23 (12%)                      | 5 (3%)                      | $\chi^2=10.15 p=.017$  |
| Witness not Wearing His/Her Glasses | 7 (4%)                             | 10 (6%)                       | 6 (3%)                        | 17 (11%)                    | $\chi^2=10.87 p=.012$  |
| Location                            |                                    |                               |                               |                             | $\chi^2= 52.96 p=.000$ |
| Police Facility <sup>a</sup>        | 46 (24%)                           | 34 (20%)                      | 86 (46%)                      | 51 (32%)                    |                        |
| Witness Home                        | 50 (26%)                           | 61 (35%)                      | 56 (30%)                      | 48 (30%)                    |                        |
| Witness Work/school                 | 48 (25%)                           | 46 (26%)                      | 30 (16%)                      | 31 (19%)                    |                        |
| Public Setting                      | 40 (21%)                           | 31 (18%)                      | 11 (6%)                       | 26 (16%)                    |                        |
| Other                               | 10 (5%)                            | 3 (2%)                        | 4 (2%)                        | 6 (4%)                      |                        |
| Serial Investigation                | 80 (41%)                           | 77 (44%)                      | 76 (41%)                      | 50 (31%)                    | $\chi^2=6.88 p=.076$   |
| Victim                              | 166 (86%)                          | 157 (90%)                     | 166 (89%)                     | 145 (90%)                   | $\chi^2=2.01 p=.570$   |
| Suspect in Position #1 <sup>b</sup> | 28 (14%)                           | 19 (11%)                      | 18 (10%)                      | 13 (8%)                     | $\chi^2=4.20 p=.241$   |
| Good Viewing Opportunity            | 65 (34%)                           | 68 (39%)                      | 62 (33%)                      | 55 (34%)                    | $\chi^2=1.69 p=.640$   |
| Weapon Used                         | 136 (70%)                          | 118 (67%)                     | 123 (66%)                     | 110 (68%)                   | $\chi^2=0.83 p=.841$   |

<sup>a</sup> This category does not include “police vehicles,” these are coded “other.”

<sup>b</sup> Suspect position #1 is examined because this position is underutilized compared to other positions (11% of the time the suspect is in position #1).

### Analysis #1 Results

Measuring the outcomes of photo spread and lineup identification procedures is not necessarily a straightforward task because there are multiple ways of combining and separating the categories listed in Table 2. For example, one simple way is to measure whether the witness selected any person from the photo spread. This measure allows for an understanding of whether one procedure reduces or increases the chances that any photo will be selected. A next logical step is to examine *only* cases in which a photo was selected and determine if the four groups differ in terms of how often a suspect or filler is selected. The use of different confidence categories by HPD investigators means there are several ways to measure suspect and filler selections. One possibility is to treat positive selections as different from tentative selections; tentative selections could be considered different types of selections, such as “unsure” or “maybe.” Another possibility is to combine all three categories together and treat all identifications (positive, strong tentative, and weak tentative) as being the same. The results discussed below utilize different methods of combining and separating selection categories in order to fully explore whether photo spread identification procedures are associated with different selection results.

Table 4 examines differences between the four photo spread administration methods in terms of whether any member of the photo spread was selected (suspect and filler selections are combined together). Blinded simultaneous procedures are most likely to produce “no selection” outcomes; the difference between the blinded simultaneous method and the blinded sequential method is statistically significant. The differences between the other photo spread administration methods are not statistically different from zero.

Table 4. Differences in selecting a photo across experimental treatments.

| Selection Result   | Photo Spread Administration Method |                    |                    |                  | Total |
|--------------------|------------------------------------|--------------------|--------------------|------------------|-------|
|                    | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |       |
| No Selection       | 96 (50%)                           | 54 (31%)           | 71 (38%)           | 59 (36%)         | 280   |
| Photo was Selected | 98 (51%)                           | 121 (69%)          | 116 (62%)          | 103 (63%)        | 438   |
| <i>Total</i>       | 194                                | 175                | 187                | 162              | 718   |

*Model significance:  $\chi^2=14.38$ ,  $p=.002$ ; a z-test is used to compare column proportions*

The results in Table 3 suggest there are important differences between the four photo spread groups in terms of some variables. For this reason a multivariate logistic regression model was estimated to control for the variables identified as being important: corroborating evidence, the use of an interpreter, witness use of drugs or alcohol at the time of the robbery, the witness did not wear his/her glasses during the robbery, the witness saw a photo of the perpetrator between the time of the crime and the photo spread procedure, the witness was shown the photo spread at a police facility, and the witness was shown a photo spread at his/her workplace or school. The multivariate results are presented in Appendix C (Table C-1) and confirm the bivariate results that show the blinded simultaneous method (i.e., all photos are shown at the same time on a single sheet of paper by an investigator who knows the identity of the suspect but does not know the suspect’s position in the photo spread) reduces the chances that a photo was selected compared to the other three methods. The other three methods do not differ from one another.

The next step is to separate the “photo selected” category into filler and suspect selections. For this analysis the selection result is measured with three categories: no selection, suspect selection (positive, strong tentative, and weak tentative combined together), and filler selection (positive, strong tentative, and weak tentative combined together). Table 5 shows “no selection” results are most common among the blinded simultaneous group, which is statistically different from the blinded sequential group (50% compared to 31%). The blinded simultaneous method produced the lowest rate of suspect identifications (25%) while the blinded sequential method produced the highest rate of suspect identifications (39%). The difference between these two methods is statistically significant. Filler selection rates are roughly the same across the four methods and the differences are not statistically significant. No method appears to perform better in terms of reducing filler selections.

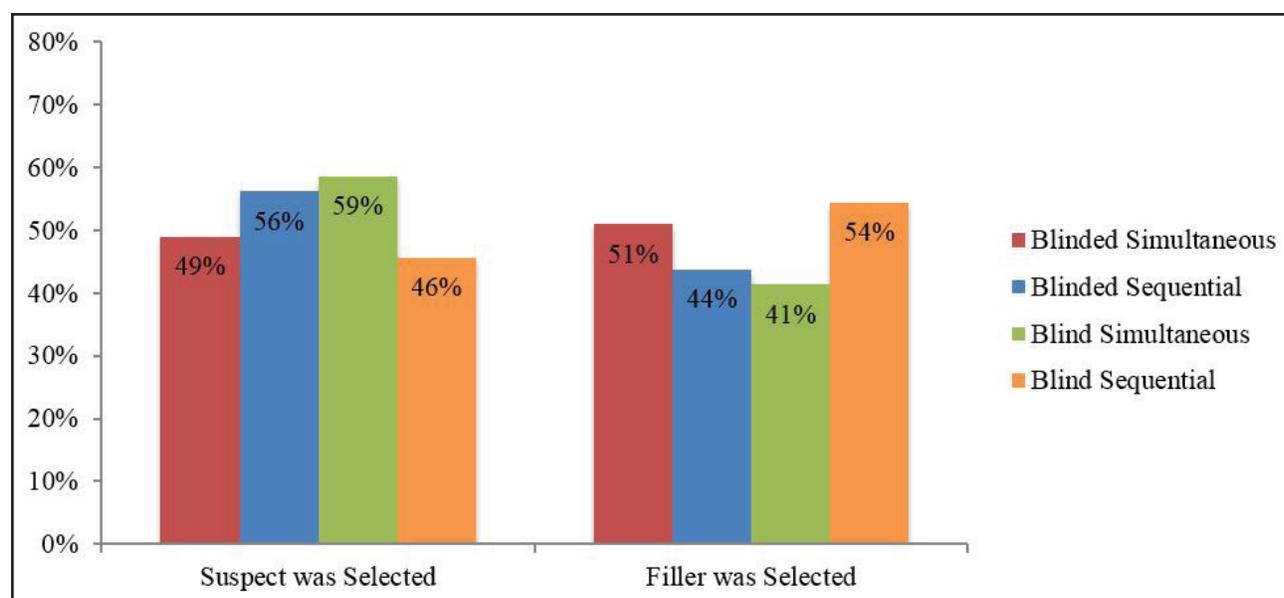
Table 5. Differences in selecting a photo, a filler, and the suspect across experimental treatments.

| Selection Result     | Photo Spread Administration Method |                    |                    |                  | Total |
|----------------------|------------------------------------|--------------------|--------------------|------------------|-------|
|                      | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |       |
| No Selection         | 96 (50%)                           | 54 (31%)           | 71 (38%)           | 59 (36%)         | 280   |
| Suspect was Selected | 48 (25%)                           | 68 (39%)           | 68 (36%)           | 47 (29%)         | 231   |
| Filler was Selected  | 50 (26%)                           | 53 (30%)           | 48 (26%)           | 56 (35%)         | 207   |
| Total                | 194                                | 175                | 187                | 162              | 718   |

Model significance:  $\chi^2=19.31$ ,  $p=.004$ ; a z-test is used to compare column proportions

Figure 1 presents the percentages of suspect and filler selection results across the four photo spread methods. The percentages in Figure 1 are based only on cases in which a selection occurred and reflect the patterns in Table 5. Filler selections (i.e., mistaken identifications) occur more frequently than suspect selections when blinded simultaneous and blind sequential methods are used. Suspect selections occur more frequently when blinded sequential and blind simultaneous methods are used. Despite the apparent differences between the methods (i.e., blind simultaneous compared to blind sequential), these differences are not statistically significant. In other words, the differences between the photo spread methods are not (statistically) different from zero.

Figure 1. Suspect and filler selections among cases in which a selection was made.



As noted above, it is important to compare the four photo spread administration methods in a multivariate model to control for the variables identified in Table 3 as distinguishing the treatment groups: corroborating evidence, the use of an interpreter, witness use of drugs or alcohol at the time of the robbery, the witness did not wear his/her glasses during the robbery, the witness saw a photo of the perpetrator between the time of the crime and the photo spread procedure, the witness was shown the photo spread at a police facility, and the witness was shown a photo spread at his/her

workplace or school. The multivariate results are presented in Appendix C (Table C-2) and, overall, mirror the bivariate results in Table 5.

The measure of witness selections can be divided into finer categories by disaggregating suspect and filler picks into “positive” selections and “tentative” selections. This allows for an understanding of whether and how photo spread administration methods may differ in terms of the “certainty” of selections. Table 6 shows the comparison of the photo administration methods in terms of this measure.

Table 6. Differences in selecting a photo, a filler, and the suspect across experimental treatments: Disaggregating positive from tentative identifications.

| Selection Result     | Photo Spread Administration Method |                    |                    |                  | Total      |
|----------------------|------------------------------------|--------------------|--------------------|------------------|------------|
|                      | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |            |
| No Selection         | 96 (50%)                           | 54 (31%)           | 71 (38%)           | 59 (36%)         | 280        |
| Positive Suspect ID  | 30 (16%)                           | 34 (19%)           | 49 (26%)           | 23 (14%)         | 136        |
| Tentative Suspect ID | 18 (9%)                            | 34 (19%)           | 19 (10%)           | 24 (15%)         | 95         |
| Positive Filler ID   | 7 (4%)                             | 6 (3%)             | 9 (5%)             | 8 (5%)           | 30         |
| Tentative Filler ID  | 43 (22%)                           | 47 (27%)           | 39 (21%)           | 48 (30%)         | 177        |
| <b>Total</b>         | <b>194</b>                         | <b>175</b>         | <b>187</b>         | <b>162</b>       | <b>718</b> |

Model significance:  $\chi^2=30.518, p=.002$ ; a z-test is used to compare column proportions

Table 6 shows that positive filler selections are rare events (n = 30) and are evenly distributed across the four administration methods. Positive suspect identifications are also fairly uncommon, occurring about 19% of the time. Blind simultaneous showings produce the highest rate of positive suspect selections (26%), followed by blinded sequential (19%), blinded simultaneous (16%), and blind sequential (14%). The difference between blind simultaneous and blind sequential in terms of suspect selections is statistically significant; positive suspect selections occur more often when blind simultaneous is used compared to blind sequential. Tentative suspect selections were significantly lower among blinded simultaneous showings than blinded sequential showings. The rates of filler selections are not statistically different across the four methods, similar to what is presented in Table 5.

The low rate of positive filler identifications makes it difficult to estimate more sophisticated multivariate models discussed above and in Appendix C. To overcome this limitation the two filler identification categories (positive and tentative) were combined together and the analysis reported in Table 6 was replicated. The results obtained when the two filler identification categories were combined are similar to those reported in Table 6.

A multivariate model was estimated using the modified selection measure with four categories: no selection, positive suspect identification, tentative suspect identification, and filler identification (see Table C-3 in Appendix C). The pattern of results mirror those reported in Table 6. In relation to blinded simultaneous showings, blind simultaneous showings increase the chances of a positive suspect selection compared to no selection at all. In relation to blinded simultaneous showings, both sequential methods increase the chances of a tentative suspect selection and a filler selection

compared to no selection at all. Finally, the blind simultaneous procedure increases the chances of a positive suspect selection rather than a filler selection compared to the blind sequential procedure.

### Analysis #1 Summary

The results presented in Tables 4 through 6 and in Appendix C show that the selection results of the four photo spread methods are similar; no one method stands out as performing substantially different from the others. The blinded simultaneous method (i.e., photos are shown at the same time on a single sheet of paper by an investigator who knows the identity of the suspect but does not know the suspect's position in the photo spread) increases the chances that the witness will reject the lineup (i.e., not make a selection). Witnesses rejected half of the blinded simultaneous photo spreads whereas rejections occurred less frequently (less than 40%) when other methods were used (see Table 4). Filler selections were about evenly distributed across the four methods (see Table 5). Suspect selections, combining positive and tentative identifications, were least likely when the blinded simultaneous method was used (see Table 5). Positive suspect identifications were more likely when the blind simultaneous method was used (see Table 6). Results of the bivariate analyses are confirmed by the results of multivariate models that control for differences between the groups identified in Table 3 (see Tables C-1 through C-3 in Appendix C).

### Analysis #2 Patterns of results among 830 photo spreads

The second set of results is based on the analysis of a different sub-sample of photo spreads (see Appendix A). This sub-sample includes robberies involving strangers and witnesses who had not previously viewed a photo spread with this suspect. This analysis *includes* photo spreads that did not follow the experimental protocol that should have been used during the investigation. From this set of 832 cases, 2 were excluded because it was not clear a blinding method was used (see Table A-3). This produced a sample of 830 for Analysis #2. The purpose of this analysis is to determine if the results from Analysis #1 are sensitive to the inclusion and exclusion of cases that violated the experimental protocol. All cases in Analysis #2 are placed in the experimental groups based on the method that was used to administer the photo spread. For example, if a photo spread was assigned to be shown blind simultaneous but was shown blinded simultaneous instead, this case is placed in the blinded simultaneous category. Analysis #2 follows the approach used in Analysis #1 above. The first step compares the treatment groups in terms of relevant variables contained in the investigator survey.

Table 7 presents the comparison of the four treatment groups in terms of relevant variables. Similar to the results in Table 3, the experimental groups differ in some ways. The patterns in Tables 3 and 7 are similar. There are differences between the treatment groups in terms of interpreters used, the presence of corroborating evidence, witnesses who were under the influence of drugs or alcohol at the time of the robbery, witnesses reported to have seen a photo of the perpetrator between the crime and the time of the identification procedure, witnesses who did not wear their glasses at the time of the robbery, the location of the photo spread administration procedure, and serial investigations. Photo spreads conducted as part of serial investigations approached, but did not attain statistical significance in Analysis #1 (see Table 3), but did reach statistical significance in this sample of 830 cases. Differences between the treatment groups are controlled during the multivariate analyses of witness selection results.

Table 7. Differences and similarities between treatment groups: Sample of 830 cases.

| Case Characteristics                | Photo Spread Administration Method |                    |                    |                  | Statistical Test       |
|-------------------------------------|------------------------------------|--------------------|--------------------|------------------|------------------------|
|                                     | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |                        |
|                                     | N = 243                            | N = 180            | N = 224            | N = 183          |                        |
| Interpreter Used                    | 6 (3%)                             | 13 (7%)            | 18 (8%)            | 27 (15%)         | $\chi^2=22.23 p=.000$  |
| Corroborating Evidence              | 158 (65%)                          | 163 (91%)          | 156 (70%)          | 121 (66%)        | $\chi^2=40.41 p=.000$  |
| Serial Investigation                | 124 (51%)                          | 79 (44%)           | 102 (46%)          | 60 (33%)         | $\chi^2=14.43 p=.002$  |
| Witness Under Influence             | 17 (7%)                            | 14 (8%)            | 3 (1%)             | 14 (8%)          | $\chi^2=11.26 p=.010$  |
| Witness Saw Photo                   | 30 (12%)                           | 16 (9%)            | 26 (12%)           | 7 (4%)           | $\chi^2=10.36 p=.016$  |
| Witness not Wearing His/Her Glasses | 8 (3%)                             | 10 (6%)            | 6 (3%)             | 19 (10%)         | $\chi^2=14.75 p=.002$  |
| Location                            |                                    |                    |                    |                  | $\chi^2= 58.02 p=.000$ |
| Police Facility <sup>a</sup>        | 51 (21%)                           | 35 (19%)           | 91 (41%)           | 61 (33%)         |                        |
| Witness Home                        | 54 (22%)                           | 63 (35%)           | 65 (29%)           | 52 (29%)         |                        |
| Witness Work/school                 | 61 (25%)                           | 44 (24%)           | 40 (18%)           | 30 (16%)         |                        |
| Public Setting                      | 51 (21%)                           | 29 (16%)           | 19 (9%)            | 25 (14%)         |                        |
| Other                               | 26 (11%)                           | 9 (5%)             | 9 (4%)             | 15 (8%)          |                        |
| Victim                              | 212 (87%)                          | 162 (90%)          | 199 (89%)          | 165 (90%)        | $\chi^2=1.20 p=.754$   |
| Suspect in Position #1 <sup>b</sup> | 36 (15%)                           | 20 (11%)           | 22 (10%)           | 14 (8%)          | $\chi^2=5.98 p=.112$   |
| Good Viewing Opportunity            | 76 (31%)                           | 70 (39%)           | 69 (31%)           | 60 (33%)         | $\chi^2=3.63 p=.304$   |
| Weapon Used                         | 181 (75%)                          | 120 (67%)          | 153 (68%)          | 119 (65%)        | $\chi^2=5.28 p=.152$   |

<sup>a</sup> This category does not include “police vehicles,” these are coded “other.”

<sup>b</sup> Suspect position #1 is examined because this position is underutilized compared to other positions (11% of the time the suspect is in position #1).

### Analysis #2 Results

Table 8 displays similarities and differences between the four photo spread administration methods in terms of whether any member of the photo spread was selected (suspect and filler selections are combined together). The two simultaneous procedures are more likely to produce “no selection” outcomes than the two sequential methods; the difference between blinded simultaneous and blinded sequential is statistically significant. The pattern of results in Table 8, with the sample that includes experimental protocol violation cases, is nearly identical to the pattern in Table 4.

Table 8. Differences in selecting a photo across experimental treatments.

| Selection Result   | Photo Spread Administration Method |                    |                    |                  | Total      |
|--------------------|------------------------------------|--------------------|--------------------|------------------|------------|
|                    | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |            |
| No Selection       | 116 (48%)                          | 55 (31%)           | 92 (41%)           | 65 (35%)         | 328        |
| Photo was Selected | 127 (52%)                          | 125 (69%)          | 132 (59%)          | 118 (65%)        | 502        |
| <i>Total</i>       | <i>243</i>                         | <i>180</i>         | <i>224</i>         | <i>183</i>       | <i>830</i> |

*Model significance:  $\chi^2=14.37$ ,  $p=.002$ ; a z-test is used to compare column proportions*

A multivariate logistic regression model was estimated to control for variables identified in Table 7 as being important. The multivariate results presented in Appendix C (Table C-4) confirm the bivariate results. When the two simultaneous methods are compared there are no differences in terms of the chances a photo was selected. Similarly when the two sequential methods are compared there are no differences in terms of the chances a photo was selected. The two simultaneous procedures are less likely to produce photo selections than the two sequential procedures.

Table 9 presents the results when the selection outcomes are divided into three categories: no selection, suspect selection, and filler selection. Across the four methods, the blinded sequential method has the highest rate of suspect selections. This is similar to the pattern reported in Table 5. The simultaneous methods have similar filler selection rates (24%) that are lower than the sequential methods. The blind sequential method has the highest rate of filler selections (35%). The only difference that is statistically significant is the “no selection” rate between blinded simultaneous (48%) and blinded sequential (31%). The pattern of results in Table 9 closely mirrors Table 5.

Table 9. Differences in selecting a photo, a filler, and the suspect across experimental treatments.

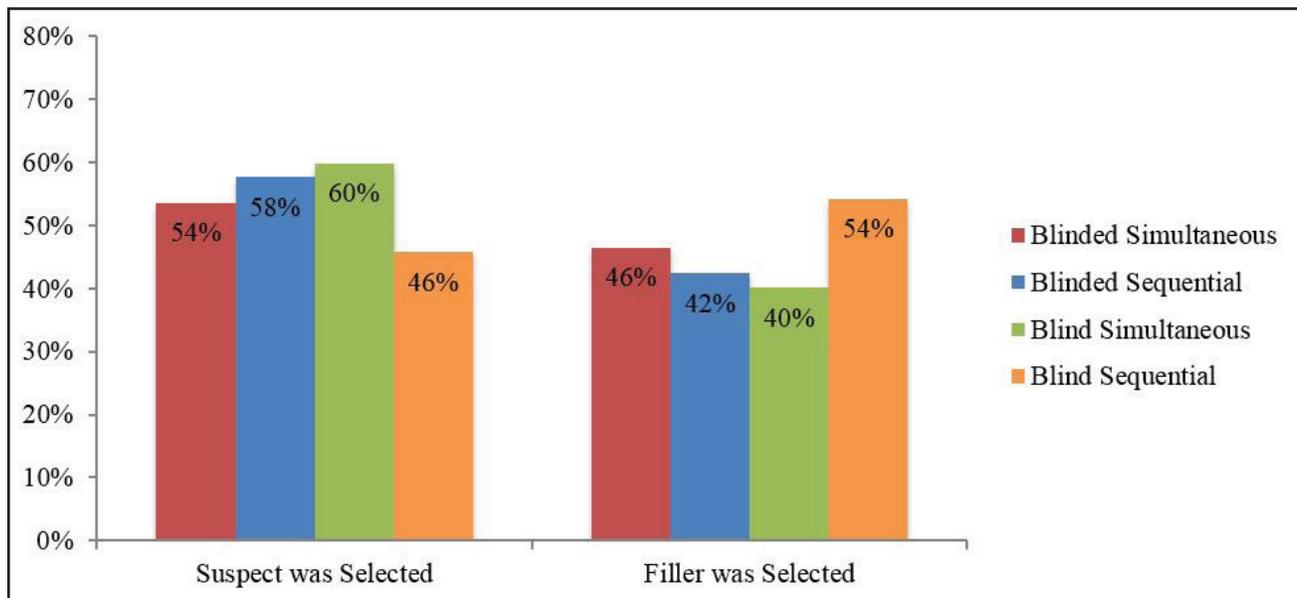
| Selection Result     | Photo Spread Administration Method |                    |                    |                  | Total      |
|----------------------|------------------------------------|--------------------|--------------------|------------------|------------|
|                      | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |            |
| No Selection         | 116 (48%)                          | 55 (31%)           | 92 (41%)           | 65 (36%)         | 328        |
| Suspect was Selected | 68 (28%)                           | 72 (40%)           | 79 (35%)           | 54 (30%)         | 273        |
| Filler was Selected  | 59 (24%)                           | 53 (29%)           | 53 (24%)           | 64 (35%)         | 229        |
| <i>Total</i>         | <i>243</i>                         | <i>180</i>         | <i>224</i>         | <i>183</i>       | <i>830</i> |

*Model significance:  $\chi^2 =20.314$ ,  $p =.002$ ; a z-test is used to compare column proportions*

A multivariate model was estimated to compare the four photo spread procedures in terms of the three-category selection variable, while controlling for relevant variables that differ across the procedures. The results of the multivariate analysis are presented in Appendix C (Table C-5) and reflect the patterns shown in Table 9. Table C-5 shows no statistical differences between the methods when filler and suspect selections are contrasted. The sequential methods are more likely to produce filler selections than “no selections” when compared to simultaneous methods. There are no clear patterns in Table 9 and Table C-5 to suggest one method or set of methods consistently protects against filler identifications.

Figure 2 presents the percentages of suspect and filler selection results across the four photo spread methods. The percentages are based on cases in which a selection occurred and reflect the patterns in Table 9. Filler selections occur more frequently than suspect selections when the blind sequential method is used. Suspect selections occur more frequently than filler selections when the other three methods are used. In Figure 2, none of the differences between the methods (i.e., blind simultaneous compared to blind sequential) is statistically significant. In other words, the differences between the photo spread methods are not (statistically) different from zero. These patterns are nearly identical to the patterns based on the analysis only of cases that followed the experimental protocol (see Figure 1).

Figure 2. Suspect and filler selections among cases in which a selection was made: Sample of 830 cases.



The final stage of Analysis #2 disaggregates filler and suspect selections into positive and tentative categories. Table 10 presents the rates of selections across the five selection categories: no selection, positive suspect selection, tentative suspect selection (weak and strong combined), positive filler identification, and tentative filler identification (weak and strong combined). Table 10 shows that blind simultaneous showings produce the highest rate of positive suspect selections (25%), followed by blinded sequential (19%), blinded simultaneous (18%), and blind sequential (16%). Blinded sequential procedures generate the highest rate of tentative suspect selections (21%), followed by blind sequential. Positive filler selections are relatively rare events (n = 33) and are evenly distributed across the four administration methods. Tentative suspect selections are more common among the two sequential methods than the two simultaneous methods. Two statistically significant differences emerged from Table 10. As expected based on the patterns described above, the blinded simultaneous method produced a greater rate of “no selection” results than the blinded sequential method. Second, both simultaneous methods produced lower rates of “tentative suspect” selections than the blinded sequential method. The patterns in Table 10 are similar to those in Table 6, implying that results are not sensitive to inclusion or exclusion of photo spread showings that deviated from the experimental conditions.

Table 10. Differences in selecting a photo, a filler, and the suspect across experimental treatments: Disaggregating positive from tentative identifications.

| Selection Result     | Photo Spread Administration Method |                    |                    |                  | Total      |
|----------------------|------------------------------------|--------------------|--------------------|------------------|------------|
|                      | Blinded Simultaneous               | Blinded Sequential | Blind Simultaneous | Blind Sequential |            |
| No Selection         | 116 (48%)                          | 55 (31%)           | 92 (41%)           | 65 (36%)         | 328        |
| Positive Suspect ID  | 43 (18%)                           | 35 (19%)           | 56 (25%)           | 29 (16%)         | 163        |
| Tentative Suspect ID | 25 (10%)                           | 37 (21%)           | 23 (10%)           | 25 (14%)         | 110        |
| Positive Filler ID   | 7 (3%)                             | 6 (3%)             | 10 (5%)            | 10 (6%)          | 33         |
| Tentative Filler ID  | 52 (21%)                           | 47 (26%)           | 43 (19%)           | 54 (30%)         | 196        |
| <i>Total</i>         | <i>243</i>                         | <i>180</i>         | <i>224</i>         | <i>183</i>       | <i>830</i> |

Model significance:  $\chi^2=31.760$ ,  $p=.002$ ; a z-test is used to compare column proportions

A multivariate model was estimated to control for differences between the experimental groups, using a four-category selection measure: no selection, positive suspect identification, tentative suspect identification (strong tentative and weak tentative combined), and filler identification (positive, strong tentative, and weak tentative combined). The results are presented in Appendix C (Table C-6) and generally reflect the patterns presented in Table 10.

In relation to blinded simultaneous showings, blind simultaneous showings increase the chances of a positive suspect selection compared to no selection at all. In relation to the simultaneous methods, the sequential procedures increase the chances of a filler selection compared to no selection at all and sequential procedures also increase the chances of a tentative suspect selection compared to no selection. Finally, the blind simultaneous procedure increases the chances of a positive suspect selection rather than a filler selection compared to the blind sequential procedure and the blinded sequential procedure.

## Analysis #2 Summary

Analysis #2 was carried out to determine if the results presented in Analysis #1 are sensitive to the inclusion or exclusion of photo spread showings that deviated from the experimental protocols. Overall, the results of Analysis #1 and #2 are similar, implying that results are not sensitive to the inclusion or exclusion of these cases. The results of multivariate models that control for differences between the experimental groups mirror the results of bivariate analyses.

## CONCLUSION

The HPD Robbery Division carried out an important experiment that provides sound evidence about the performance of four different methods of administering photo spreads to witnesses. In the current study performance was measured in terms of witness selection results. The main research question is whether four different methods of administering photo spreads to witnesses produce different selection results, such as no selection, suspect selection, and filler selection. The pattern of results show that no one method consistently performs better than the others in terms of reducing filler selections. Similarly, the methods do not appear to perform appreciably different in terms of suspect selection rates. The results reported here suggest that decisions about which photo spread procedure(s) to adopt will not be based solely on evidence about filler and suspect selection rates. HPD decision makers will need to consider other factors when deciding which method(s) to adopt. Data analysis will continue in order to explore additional research questions.

There is a relatively large body of research evidence that has been interpreted as showing a sequential method advantage (see Steblay, Dysart, & Wells, 2011; Wells et al., 1998). The results reported here are not consistent with those results. Recent views on the existing body of research evidence, however, cast doubt on the sequential method advantage (Clark, 2012; Gronlund, Carlson, Dailey, & Goodsell, 2009; Mickes, Flowe, & Wixted, 2012) and a recently completed field experiment (Wells et al., in press) also casts doubt on the notion of a strong sequential method advantage. The lack of differences between methods that are reported here is consistent with more recent discussions in the research literature. In addition, research on the sources of erroneous convictions demonstrates that reforms should be multi-faceted and should not focus exclusively on police eyewitness identification procedures (Gould, Carrano, Leo, & Young, 2013).

One important lesson for HPD and other law enforcement agencies is that complex field studies that assess photo spread administration methods are feasible. Law enforcement agencies can expand the current state of knowledge about eyewitness identification methods by using their existing data and collecting new data. Additional field studies are necessary to complement the large number of studies that have been conducted in laboratory settings and to provide new, systematic information about eyewitness identifications that occur as part of actual criminal investigations (see IACP, 2013). The growth of information derived from field studies will enhance the policy and procedural recommendations that follow from this body of research.

Successful field studies do not need to use complex research designs, like the one described here, in order to provide valuable information that other practitioners and researchers can use. There are many questions about eyewitness identifications and identification procedures that have not been examined with field data. Reporting basic descriptive information about eyewitness identifications and reporting data about simple bivariate relationships between meaningful variables will continue to expand research knowledge in positive ways. Field studies are not easy to implement but can be successfully carried out when individuals across different ranks in an agency are committed to conducting a high-quality study. Researcher – practitioner partnerships can also play a role in conducting sound research.

## REFERENCES

- Clark, S.E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science, 7*, 238-259.
- Doob, A.N., & Kirshenbaum, H.M. (1973). Bias in police lineups: Partial remembering. *Journal of Police Science and Administration, 1*, 287-293.
- Feder, L., Jolin, A., & Feyerherm, W. (2000). Lessons from two randomized experiments in criminal justice settings. *Crime & Delinquency, 46*, 380-400.
- Gould, J., Carrano, J., Leo, R., & Young, J. (2013). *Predicting erroneous convictions: A social science approach to miscarriages of justice* (Publication No. NCJRS 241389). Retrieved from the National Institute of Justice: <http://www.nij.gov/topics/justice-system/wrongful-convictions/Pages/predicting-preventing.aspx>.
- Gronlund, S.D., Carlson, C.A., Dailey, S.B., & Goodsell, C.A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 15*, 140-152.
- International Association of Chiefs of Police. (2013). *National Summit on Wrongful Convictions: Building a systematic approach to prevent wrongful convictions*. Alexandria, VA: International Association of Chiefs of Police.
- McGarrell, E.F., Chermak, S., Weiss, A., & Wilson, J. (2001). Reducing firearm violence through directed police patrol. *Criminology & Public Policy, 1*, 119-148.
- Malpass, R.S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior, 5*, 299-309.
- Malpass, R.S., Tredoux, C.G., & McQuiston-Surrett, D.E. (2007). Lineup construction and lineup fairness. In R.Lindsay, D., Ross, J.D. Read, & M.P. Toggia (Eds.), *Handbook of eyewitness psychology, (Vol. 2): Memory for people*. Lawrence Erlbaum & Associates.
- Mazerolle, L.G., Roehl, J., & Kadleck, C. (1998). Controlling social disorder using civil remedies: Results from a randomized field experiment in Oakland, California. *Crime Prevention Studies, 9*, 141-159.
- Mickes, L., Flowe, H.D., & Wixted, J.T. (2012). Receiver operator characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18*, 361-376.
- Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth.
- Sherman, L.W., Schmidt, J.D., & Rogan, D.P. (1992). *Policing domestic violence: Experiments and dilemmas*. New York, NY: Free Press.
- Stebly, 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(1)*, 99-139.
- Weisburd, D. (2000). Randomized experiments in criminal justice policy: Prospects and problems. *Crime & Delinquency, 46*, 181-193.
- Weisburd, D., Petrosino, A., & Mason, G. (1993). Design sensitivity in criminal justice experiments. *Crime and Justice, 17*, 337-379.
- Wells, G.L., Leippe, M.R., & Ostrom, T. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior, 3*, 285-293.

- Wells, G.L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedure: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 1-39.
- Wells, G.L., Steblay, N.K., & Dysart, J.E. (In Press). Double-blind photo lineups using actual eyewitnesses: An experimental test of a sequential versus simultaneous lineup procedure. *Law and Human Behavior*.
- Wells, G.L., Steblay, N.K., & Dysart, J.E. (2011). *A test of the simultaneous vs. sequential lineup methods: An initial report of the AJS national eyewitness identification field studies*. Des Moines, IA: American Judicature Society.
- Wells, W. (2013). *The Houston Police Department eyewitness identification experiment: Sampling, research design, and data collection: A report to the Houston Police Department*. Unpublished manuscript, Sam Houston State University.

## APPENDIX A

The following steps were used to identify the sample of cases included in the main analysis.

Step 1: 997 photo spreads (see Table 1).

Step 2: 855 photo spreads with witnesses and perpetrators who were strangers. 142 cases were excluded between step 1 and step 2.

Step 3: 845 photo spreads involving strangers and the suspect was not previously shown to the witness in a photo spread. 10 cases were excluded between step 2 and step 3.

Step 4: 832 photo spreads involving strangers and the suspect was not previously shown to the witness in a photo spread and data are available about the final selection result (i.e., the result is not ambiguous as indicated in Table 2). 13 cases were excluded between step 3 and step 4.<sup>9</sup>

Step 5: 718 cases photo spreads involving strangers and the suspect was not previously shown to the witness in a photo spread and data are available about the final selection result and the experimental protocol was followed. 114 cases were excluded between step 4 and step 5.

The exclusion of cases because the experimental protocols were not followed are of concern because when these cases are excluded it could lead to groups that are systematically different from one another. If the groups are systematically different then there are a number of reasons why the groups may differ in terms of witness identification results. For this reason, the tables listed below presents information about the patterns of protocol violations.

Table A-1 presents the percentage of cases within each experimental treatment group that involve a protocol violation. This analysis examines the 832 photo spread cases identified in Step 4 described above. These cases are photo spreads involving witnesses and perpetrators who are strangers, cases in which the witness had not previously viewed a suspect in an identification procedure, and cases that were not missing data on important variables. Table A-1 shows the distribution of protocol violations across the methods that should have been used during the identification procedure. Overall, protocols were followed 86% of the time (718/832). Within the four experimental groups, protocols were followed between 81% and 92% of the time.<sup>10</sup> Protocol violations were more common when a sequential procedure (i.e., photos are shown individually, one at a time, to the witness) should have been used. When a blinded simultaneous method procedure (i.e., all photos are shown at the same time on a single sheet of paper by an investigator who knows the identity of the suspect but does not know the suspect's position in the photo spread) was to be used the protocol was followed in 91% of the cases. Similarly, when a blind simultaneous method (i.e., all photos are shown at the same time on a single sheet of paper by an investigator who does not know the identity of the suspect and does not know the suspect's position in the photo spread) was to be used the protocol was also followed in 92% of the cases. When a blinded sequential method (i.e., photos are shown individually, one at a time, by an investigator who knows the identity of the suspect but does not know the suspect's position in the photo spread) was to be used the protocol was followed in 81% of the cases. When a blind sequential method (i.e., photos are shown individually, one at a time, by an investigator who does not know the identity of the suspect and does not know the suspect's position in the photo spread) was to be used the protocol was followed in 82% of the cases.

<sup>9</sup> Table 2 shows there were 18 cases with selection results coded as "other." Step 4 shows 13 cases were excluded rather than 18 because 5 had already been excluded in previous steps (1 of the 18 was a video lineup, 2 were video lineups, and 2 were non-stranger cases).

<sup>10</sup> This range of protocol adherence is similar to that obtained in the Minneapolis Domestic Violence Experiment and four replication cities (Sherman et al., 1992, p. 15).

Table A-1. Identification procedures and experimental protocol violations.

|                    | Method Assigned             |                           |                           |                         | <i>Total</i> |
|--------------------|-----------------------------|---------------------------|---------------------------|-------------------------|--------------|
|                    | <b>Blinded Simultaneous</b> | <b>Blinded Sequential</b> | <b>Blind Simultaneous</b> | <b>Blind Sequential</b> |              |
| Protocol Violation | 19 (9%)                     | 42 (19%)                  | 17 (8%)                   | 36 (18%)                | 114          |
| No Violation       | 194 (91%)                   | 175 (81%)                 | 187 (92%)                 | 162 (82%)               | 718          |
| <i>Total</i>       | <i>213</i>                  | <i>217</i>                | <i>204</i>                | <i>198</i>              | <i>832</i>   |

Table A-2 presents more detailed information about the nature of the protocol violations. This table shows that the random assignment method used during the study produced similar numbers of cases within each assigned treatment group. A blinded simultaneous method was to be used in 213 cases, a blinded sequential procedure was to be used in 217 cases, blind simultaneous in 204 cases, and blind sequential in 198 cases.

Table A-2. Identification procedures assigned and actually used.

| <b>Method Used</b>   | Method Assigned              |                           |                           |                         | <i>Total</i> |
|----------------------|------------------------------|---------------------------|---------------------------|-------------------------|--------------|
|                      | <b>Blinded Simultaneous</b>  | <b>Blinded Sequential</b> | <b>Blind Simultaneous</b> | <b>Blind Sequential</b> |              |
| Blinded Simultaneous | <b>196 (92%)<sup>a</sup></b> | 24 (11%)                  | 13 (6%)                   | 12 (6%)                 | 245          |
| Blinded Sequential   | 0 (0%)                       | <b>175 (81%)</b>          | 0 (0%)                    | 5 (3%)                  | 180          |
| Blind Simultaneous   | 13 (6%)                      | 5 (2%)                    | <b>187 (92%)</b>          | 19 (10%)                | 224          |
| Blind Sequential     | 4 (2%)                       | 13 (6%)                   | 4 (2%)                    | <b>162 (82%)</b>        | 183          |
| <i>Total</i>         | <i>213</i>                   | <i>217</i>                | <i>204</i>                | <i>198</i>              | <i>832</i>   |

<sup>a</sup> Two of these cases involve a violation of the photo spread preparation method (i.e., the showing investigator also shuffled the spread). These two cases are counted as protocol violations in Table A-1 above and are also included in the counts of protocol violations in Table A-3 below.

Finally, Table A-3 presents information about the set of 114 protocol violations. The largest number of violations occurred in cases when a blinded sequential method should have been used (n = 42) and when a blind sequential method should have been used (n = 36), followed by blinded simultaneous (n = 19) and blind simultaneous (n = 17). These experimental protocol violations were driven, primarily, by photo spread showings that occurred as part of serial robbery investigations. Out of the 114 cases in which a protocol violation occurred, 72% (n = 82) of the procedures were conducted as part of a serial investigation. More specifically, 5 different serial robbery investigations generated 36% (41/114) of the experimental protocol violation cases. It was common for serial investigations to involve multiple photo spread showings. Thus, experimental violations that occurred during a serial investigation generated several instances of a protocol violation. Finally, 44% (n = 36) of the experimental protocol violation showings that were conducted as part of a serial investigation actually followed the protocol that would have been assigned to the case if it was not part of a serial investigation (see Wells, 2013 for a discussion of the random assignment procedures). The patterns in

Tables A-1, A-2, and A-3 show it is important to compare the four treatment groups in terms of relevant variables to understand the degree to which the groups are roughly equivalent.

Table A-3. Experimental treatment that should have been used and treatments actually used.

| <b>Method Used</b>   | <b>Method Assigned</b>      |                           |                           |                         | <b>Total</b> |
|----------------------|-----------------------------|---------------------------|---------------------------|-------------------------|--------------|
|                      | <b>Blinded Simultaneous</b> | <b>Blinded Sequential</b> | <b>Blind Simultaneous</b> | <b>Blind Sequential</b> |              |
| Blinded Simultaneous | 2 <sup>a</sup>              | 24 <sup>b</sup>           | 13 <sup>c</sup>           | 12 <sup>d</sup>         | 51           |
| Blinded Sequential   | 0                           | 0                         | 0                         | 5                       | 5            |
| Blind Simultaneous   | 13                          | 5                         | 0                         | 19 <sup>e</sup>         | 37           |
| Blind Sequential     | 4                           | 13                        | 4                         | 0                       | 21           |
| <i>Total</i>         | <i>19</i>                   | <i>42</i>                 | <i>17</i>                 | <i>36</i>               | <i>114</i>   |

<sup>a</sup> These two cases involve a violation of the photo spread preparation method (i.e., the showing investigator also shuffled the spread).

<sup>b</sup> Nine of these twenty-four photo spread showings occurred as part of one serial robbery investigation.

<sup>c</sup> Ten of these thirteen photo spread showings occurred as part of two serial robbery investigations.

<sup>d</sup> All twelve photo spread showings occurred as part of one serial robbery investigation.

<sup>e</sup> Fourteen of these nineteen photo spread showings occurred as part of two different serial robbery investigations.

## APPENDIX B

### Lineup Fairness Assessment

Stryker Swindle, M.A.<sup>11</sup>

Assessing the quality (i.e., fairness and bias) of photo spreads included in the experiment is important because the quality of photo spreads and lineups should be roughly equivalent across all four experimental groups. To examine the potential differences between the four experimental groups the fairness of photo spreads was examined for a sample of 60 photo spreads. Fifteen photo spreads were randomly selected from each of the four experimental conditions. The sample of mock witnesses included 15 HPD police cadets, 15 volunteers from the HPD Positive Interaction Program (PIP), and 19 Sam Houston State University students. Each mock witness viewed the 60 photo spreads with the descriptions of the robbery perpetrator and/or suspect.

Since the early 1970's social scientists have been estimating the degree of bias in photo spreads by using mock witnesses (Doob & Kirshenbaum, 1973). Mock witnesses are individuals with no knowledge of the offense who are asked to identify the suspect from a lineup. Although methodologies differ, mock witnesses are typically given a description of the perpetrator provided by an actual eyewitness after the offense. Definitions of fairness differ, but an essential element is that the suspect does not "stand out" to mock witnesses from the other members of a photo spread. This Appendix describes the methods and findings of the fairness assessment conducted during the HPD field experiment.

### Methods

Sixty photo spreads were randomly selected from the first 500 identification procedures conducted during experiment. Fifteen photo spreads were selected from each of the four groups in the experiment: blind simultaneous, blind sequential, blinded simultaneous, and blinded sequential. Some photo spreads that were selected contained the same suspect and fillers. When duplicate photo spreads were obtained, a replacement photo spread was randomly selected so that the mock witnesses did not view a photo spread with the same lineup members more than once. In collaboration with HPD personnel, members of the SHSU research team reviewed OLO reports for each of the photo spreads and recorded the description of the perpetrator or suspect contained in the investigative report. The position of the suspect in each procedure was obtained from OLO reports with assistance from HPD personnel in order to verify the position of the suspect recorded in the survey data.

The sample of mock witnesses who viewed the photo spreads consisted of college students, police cadets from the HPD Training Academy, and civilian volunteers from the HPD PIP. College students included undergraduate and graduate students from Sam Houston State University who volunteered to participate. African American and Hispanic student organizations were approached to ensure the inclusion of diverse racial groups. A large proportion of the students included in the study were undergraduate criminal justice majors. Additionally, 15 cadets from the HPD Training Academy and 15 volunteers from the HPD PIP also participated in the study to expand the range of participants' backgrounds. Participation in the study was voluntary and participants received no financial compensation or extra class credit.

---

<sup>11</sup> Stryker Swindle is a Ph.D. student in the Department of Criminal Justice and Criminology at Sam Houston State University.

The research team developed descriptions of perpetrators from the information included in OLO reports and attached these descriptions to each photo spread. Mock witnesses were asked to read the description and then mark, on a separate answer sheet, the individual in the photo spread they believed was the perpetrator. Mock witnesses also completed a short demographic survey that measured their age, sex, and race. Participants attended group sessions with varying numbers of other mock witnesses to complete the tasks. Sessions with college students were conducted on the campus of Sam Houston State University, sessions with police cadets were held at the HPD training academy, and sessions with HPD PIP volunteers were held at the Houston Police Officers' Union building.

Researchers have used several methods of estimating photo spread bias. Many studies of photo spread fairness measure *functional size* as the indicator of bias (Wells, Leippe, & Ostrom, 1979) and *effective size* as the indicator of lineup size (Malpass, 1981). *Bias* refers to the degree in which the suspect "stands out" from other lineup members, while *size* is the number of lineup members that are realistic choices given the description of the perpetrator (Malpass, Tredoux, & McQuiston-Surrett, 2007). When functional size (the measure of bias) is lower than the number of individuals included in the lineup (i.e., 6), the procedure may be biased *toward* a suspect. In these photo spreads, the suspect "stands out" to mock witnesses and is chosen more frequently than expected by chance alone. If functional size is greater than the number of lineup members, the photo spread may be biased *away* from the suspect. When the effective size of the lineup is substantially less than the number of lineup members, the photo spread may not include many suitable alternatives to the suspect based on the description of the perpetrator given by the witness.<sup>12</sup> Because measures of bias and size differ, both measures were calculated for each of the photo spreads to provide a more comprehensive assessment of fairness.<sup>13</sup>

## Findings

Tables B-1 through B-4 present the proportion of mock witnesses who selected the suspect from each photo spread for each of the four experimental conditions. These tables present the photo spread numbers in the first row and the proportion of mock witnesses who selected the suspect from that photo spread in the second row. If the suspect is chosen by the proportion expected by chance in a six-person photo spread, the proportion of mock witnesses who select should be .167 (1/k, where *k* is the number of individuals in the photo spread). A proportion greater than .167 shows the suspect was chosen more often than what would be expected by chance and a proportion less than .167 shows the suspect was selected less often than what would be expected by chance.

The photo spread with the greatest proportion of mock witnesses who chose the suspect was photo spread number 115 in the blinded sequential condition. Of the mock witnesses who viewed this photo spread, 92% selected the suspect. With the exception of the photo spreads in the blind simultaneous condition, the proportions of mock witnesses who identified the suspect were beyond chance expectation in over half of the photo spreads.

Table B-1. Blinded simultaneous photo spreads: Percentage of mock witnesses who identified the suspect.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 271 | 433 | 519 | 126 | 124 | 328 | 389 | 518 | 373 | 81  | 110 | 186 | 82  | 87  | 414 |
| .02 | .02 | .08 | .14 | .16 | .18 | .20 | .22 | .31 | .31 | .35 | .35 | .39 | .44 | .73 |

<sup>12</sup> Tredoux's (1998) E' is recognized as the preferred measure of effective size and is utilized in this analysis.

<sup>13</sup> In the following sections, *bias* refers to functional size and *size* refers to Tredoux's E'.

Table B-2. Blinded sequential photo spreads: Proportion of mock witnesses who identified the suspect.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1   | 169 | 292 | 505 | 163 | 96  | 37  | 15  | 188 | 164 | 95  | 107 | 160 | 308 | 115 |
| .02 | .10 | .10 | .12 | .16 | .22 | .27 | .31 | .31 | .33 | .39 | .41 | .41 | .45 | .92 |

Table B-3. Blind simultaneous photo spreads: Proportion of mock witnesses who identified the suspect.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 139 | 144 | 47  | 194 | 347 | 422 | 199 | 234 | 49  | 355 | 127 | 237 | 340 | 50  | 98  |
| .06 | .08 | .08 | .08 | .10 | .10 | .12 | .14 | .16 | .16 | .16 | .20 | .22 | .38 | .43 |

Table B-4. Blind sequential photo spreads: Proportion of mock witnesses who identified the suspect.

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 322 | 419 | 65  | 365 | 459 | 77  | 249 | 135 | 443 | 17  | 286 | 21  | 290 | 22  | 280 |
| .02 | .06 | .06 | .10 | .10 | .12 | .16 | .20 | .22 | .27 | .29 | .29 | .31 | .35 | .47 |

Table B-5 presents the average values of bias (i.e., functional size) and size (i.e., effective size) for each of the four experimental conditions. The blinded sequential condition appears most “neutral;” these photo spreads did not appear to be biased toward or away from the suspect. Additionally, the blind sequential condition appeared to contain the most suitable photo spread members based on descriptions of the perpetrator. In the blind sequential condition, the average photo spread consisted of four lineup members that were reasonable choices given the description of the perpetrator.

Table B-5. Average bias and effective size of the sampled photo spreads.

|                             | <b>Bias</b> | <b>Size</b> |
|-----------------------------|-------------|-------------|
| Overall                     | 8.00        | 3.78        |
| Experimental Group          |             |             |
| <i>Blind Sequential</i>     | 8.60        | 4.02        |
| <i>Blind Simultaneous</i>   | 7.02        | 3.70        |
| <i>Blinded Sequential</i>   | 6.43        | 3.53        |
| <i>Blinded Simultaneous</i> | 9.43        | 3.86        |

The most important concern for the experiment is whether the indicators of bias and size differ across the four experimental groups. The four experimental groups were compared using analysis of variance (ANOVA) to determine whether these differences between the four photo spread methods were statistically significant. The results of separate ANOVAs conducted for bias and size are presented in Table B-6. The bias and size of photo spreads do not differ across the experimental conditions; the quality of the photo spreads are roughly equivalent.

Table B-6. One-Way ANOVA for bias and size across of experimental conditions.

|                | <b>SS</b> | <b>df</b> | <b>MS</b> | <b>F</b> | <b><i>p</i></b> |
|----------------|-----------|-----------|-----------|----------|-----------------|
| <i>Bias</i>    |           |           |           |          |                 |
| Between Groups | 26.13     | 3         | 8.71      | 0.08     | 0.97            |
| Within Groups  | 6212.83   | 55        | 112.96    |          |                 |
| Total          | 6238.97   | 58        |           |          |                 |
| <i>Size</i>    |           |           |           |          |                 |
| Between Groups | 1.99      | 3         | 0.66      | 0.63     | 0.60            |
| Within Groups  | 58.67     | 56        | 1.05      |          |                 |
| Total          | 60.65     | 59        |           |          |                 |

## APPENDIX C

### Analysis #1 Multivariate models

The first model estimates the effects of the different photo spread administration methods on the likelihood that a photo was selected. For this analysis the dependent variable has two categories (photo selected or not). Thus, a logistic regression model is estimated. The model controls for the seven variables identified in Table 3 as distinguishing the photo spread administration methods:

- corroborating evidence (yes or no)
- the use of an interpreter (yes or no)
- witness use of drugs or alcohol at the time of the robbery (yes or no)
- witness saw a photo of the perpetrator between the crime and the time of the photo spread procedure (yes or no)
- witness did not wear his/her glasses during the robbery (yes or no)
- witness was shown the photo spread at a police facility (yes or no)
- witness was shown a photo spread at his/her workplace or school (yes or no)

Table C-1 summarizes the estimated effects of the four photo spread administration methods. The findings in Table C-1 mirror those in Table 4. The likelihood of selecting any photo during the procedure is less likely when the blinded simultaneous procedure is used. The differences between the other three methods are not statistically different from zero.

Table C-1. Multivariate logistic regression results: Effects of photo spread administration method on the likelihood a photo was selected.

| Photo Spread Comparison                             | Difference   | Model Coefficient   |
|---|--|---------------------|
| Blinded sequential compared to blinded simultaneous | A selection is more likely when blinded sequential is used | B = .673, p = .004  |
| Blind simultaneous compared to blinded simultaneous | A selection is more likely when blind simultaneous is used | B = .432, p = .054  |
| Blind sequential compared to blinded simultaneous   | A selection is more likely when blind sequential is used.  | B = .757, p = .001  |
| Blinded sequential compared to blind simultaneous   | No difference  | B = .241, p = .317  |
| Blinded sequential compared to blind sequential     | No difference  | B = -.084, p = .116 |
| Blind simultaneous compared to blind sequential     | No difference  | B = -.325, p = .174 |

Model fit: -2LL = 888.429,  $\chi^2 = 71.876$ , p = .000

The second model utilizes a measure of the selection result that has three categories: no selection, suspect identification (positive and tentative selections combined), and filler identification (positive

and tentative selections combined). This three-category outcome measure requires a multinomial regression model. Presenting the results of this model is slightly more complicated than the results of the logistic regression model presented in Table C-1. This is the case because three outcome categories are being compared rather than two. Including all seven control variables generated a concern about the number of table cells with zero observations (30% of cells). The model showed that two control variables were not statistically significant: the witness was under the influence of drugs or alcohol at the time of the robbery and the witness did not wear his or her glasses at the time of the robbery. These two control variables were dropped and the model was re-estimated. This reduced the problem of cells with zero observations to 23%. Substantively, the findings regarding the effects of photo spread showing procedures were unchanged between the full and reduced models. The results described below in Table C-2 are based on the reduced model.

The pattern of results in Table C-2 are also consistent with the pattern of results in Table 5. Because any photo selection is less likely in the blinded simultaneous group, filler *and* suspect selections are more likely when blind sequential or blinded sequential methods are used. Table 5 shows filler identifications are slightly more likely when blind sequential methods are used (36%) compared to blind simultaneous methods are used (25%). The difference between these two methods in terms of whether the photo spread is rejected or a filler is selected approaches statistical significance ( $p = .090$ ) in the multivariable model (see Table C-2). The final column in Table C-2 shows the model did not detect any differences between the methods in terms of whether the witness selected a suspect or a filler. The model did not detect differences between blind and blinded sequential methods nor between blinded sequential and blind simultaneous.

Table C-2. Multivariate multinomial regression results: Effects of photo spread administration method on selection results.

| Photo Spread Comparison                             | Photo Selection Comparison   |  |                                       |
|---|--|--|---------------------------------------|
|   | No selection or filler selection   | No selection or suspect selection                                  | Filler selection or suspect selection |
| Blinded sequential compared to blinded simultaneous | A filler selection is more likely when blinded sequential is used              | A suspect selection is more likely when blinded sequential is used | No difference                         |
| Blind simultaneous compared to blinded simultaneous | No difference  | A suspect selection is more likely when blind simultaneous is used | No difference                         |
| Blind sequential compared to blinded simultaneous   | A filler selection is more likely when blind sequential is used                | A suspect selection is more likely when blind sequential is used   | No difference                         |
| Blinded sequential compared to blind simultaneous   | No difference  | No difference  | No difference                         |
| Blinded sequential compared to blind sequential     | No difference  | No difference  | No difference                         |
| Blind simultaneous compared to blind sequential     | A filler selection is more likely when blind sequential is used ( $p = .090$ ) | No difference  | No difference                         |

Model fit:  $-2LL = 318.967$ ,  $\chi^2 = 92.578$ ,  $p = .000$

The final model utilizes an outcome measure with four categories: no selection, positive suspect identification, tentative suspect identification, and filler identification. The purpose of this final model is to explore the patterns in Table 6. Positive and tentative filler identifications were combined together because of the low frequency with which fillers are positively identified (see Table 6). A multinomial, multivariate regression model was estimated including the seven control variables listed above. Including all seven control variables generated a concern about the number of table cells with zero observations (42% of cells). The model showed that two control variables could be eliminated because they were not statistically significant: the witness was under the influence of drugs or alcohol at the time of the robbery and the witness did not wear his or her glasses at the time of the robbery. These two control variables were dropped and the model was re-estimated. This reduced the problem of cells with zero observations to 33%. Substantively, the findings regarding the effects of photo spread administration procedures were unchanged between the full and reduced models.<sup>14</sup> The results described below are based on the reduced model. Table C-3 summarizes the results of this model.

The patterns in Table C-3 are consistent with the patterns presented in Table 6. Blinded simultaneous procedures are associated with lower chances of any selection. Across all comparisons, when there are differences, tentative suspect selections are more likely than positive suspect selections. The only comparison that revealed an increased chance of a positive suspect selection was between blind simultaneous and blinded simultaneous; blind simultaneous methods increased the chances of a positive suspect selection in relation to no selection at all.

Table C-3 shows the photo spread methods were compared in several ways to understand their effects on filler selections. The final two columns in the table show that the methods have very little effect on determining whether a suspect or a filler is selected. The only difference to emerge from the model in terms of filler compared to suspect selections show that blind sequential methods are associated with an increased chance of a filler selection, rather than a positive suspect identification, compared to blind simultaneous methods. The first column of results shows the blind sequential method increases the chances of a filler selection compared to no selection when compared to the blind simultaneous and blinded simultaneous methods. The blinded sequential method also increases the chances of a filler selection compared to no selection when compared to the blinded simultaneous method. Finally, no differences emerged between the blind sequential and the blinded sequential method and only one difference emerged between blind simultaneous and blinded simultaneous.

---

<sup>14</sup> One exception is that the difference between blinded simultaneous and blinded sequential in comparing tentative suspect and filler identifications attained statistical significance ( $p = .086$ ) in the reduced model but did not in the full model ( $p = .112$ )

Table C-3. Multivariate multinomial regression results: Effects of photo spread administration method on selection results, disaggregated by witness confidence.

|  | <b>Photo Selection Comparison</b>  |   |   |   |   |                                      |   |  |
|--|--|---|---|---|---|--------------------------------------|---|--|
| <b>Photo Spread Comparison</b>                         | No selection or filler selection   | No selection or positive suspect selection                                  | No selection or tentative suspect selection   | No selection or tentative suspect selection   | Positive suspect or tentative suspect selection                 | Positive suspect or filler selection | Tentative suspect or filler selection   |  |
| Blinded sequential compared to blinded to simultaneous | A filler selection is more likely when blinded sequential is used          | No difference   | A tentative suspect selection is more likely when blinded sequential is used          | A tentative suspect selection is more likely when blinded sequential is used (p=.080) | No difference   | No difference                        | A tentative suspect selection is more likely when blinded sequential is used (p=.086) |  |
| Blind simultaneous compared to blinded to simultaneous | No difference  | A positive suspect selection is more likely when blind simultaneous is used | No difference   | No difference   | No difference   | No difference                        | No difference   |  |
| Blind sequential compared to blinded to simultaneous   | A filler selection is more likely when blind sequential is used            | No difference   | A tentative suspect selection is more likely when blind sequential is used            | No difference   | No difference   | No difference                        | No difference   |  |
| Blinded sequential compared to blind simultaneous      | No difference  | No difference   | A tentative suspect selection is more likely when blinded sequential is used          | A tentative suspect selection is more likely when blinded sequential is used          | No difference   | No difference                        | No difference   |  |
| Blinded sequential compared to blind sequential        | No difference  | No difference   | No difference   | No difference   | No difference   | No difference                        | No difference   |  |
| Blind simultaneous compared to blind sequential        | A filler selection is more likely when blind sequential is used (p = .095) | No difference   | A tentative suspect selection is more likely when blind sequential is used (p = .098) | A tentative suspect selection is more likely when blind sequential is used            | A filler selection is more likely when blind sequential is used | No difference                        | No difference   |  |

Model fit: -2LL = 240.649,  $\chi^2 = 113.869$ ,  $p = .000$

### Analysis #2 Multivariate models

The first model estimates the effects of the different photo spread administration methods on the likelihood that a photo was selected. For this analysis the dependent variable has two categories (photo selected or not). Thus, a logistic regression model is estimated. Nine variables identified in Table 7 distinguish the photo spread administration methods:

- corroborating evidence (yes or no)
- the use of an interpreter (yes or no)
- witness use of drugs or alcohol at the time of the robbery (yes or no)
- witness saw a photo of the perpetrator between the crime and the time of the photo spread procedure (yes or no)
- witness did not wear his/her glasses during the robbery (yes or no)
- witness was shown the photo spread at a police facility (yes or no)
- witness was shown a photo spread at his/her workplace or school (yes or no)
- witness was shown a photo spread at his/her residence (yes or no)
- photo spread was administered as part of a serial investigation (yes or no)

Table C-4 summarizes the estimated effects of the four photo spread administration methods. The findings in Table C-4 show that sequential methods are associated with an increased chance of photo selections compared to the simultaneous methods, which is consistent with the patterns in Table 8. Table C-4 also shows there is no difference between blind simultaneous and blinded simultaneous methods nor between blind sequential and blinded sequential methods. This pattern is also consistent with the pattern in Table 8.

Table C-4. Multivariate logistic regression results: Effects of photo spread administration method on the likelihood a photo was selected.

| <b>Photo Spread Comparison</b>                      | <b>Difference</b>  | <b>Model Coefficient</b> |
|---|--|--------------------------|
| Blinded sequential compared to blinded simultaneous | A selection is more likely when blinded sequential is used | B = .660, p = .003       |
| Blind simultaneous compared to blinded simultaneous | No difference  | B = .267, p = .185       |
| Blind sequential compared to blinded simultaneous   | A selection is more likely when blind sequential is used   | B = .740, p = .001       |
| Blinded sequential compared to blind simultaneous   | A selection is more likely when blinded sequential is used | B = .393, p = .081       |
| Blinded sequential compared to blind sequential     | No difference  | B = -.079, p = .741      |
| Blind simultaneous compared to blind sequential     | A selection is more likely when blind sequential is used   | B = .473, p = .034       |

Model fit: -2LL = 1034.386,  $\chi^2 = 79.489$ , p = .000

The second model uses the three-category outcome measure: no selection, suspect selection (positive and tentative combined), and filler selection (positive and tentative combined). The inclusion of nine control variables, in addition to the variables that measure the photo spread administration procedure, may lead to problems with the model. Specifically, the inclusion of this many independent variables and a three-category dependent variable meant there were many instances (35% of cells) in which combinations of variables generated zero observations. For this reason, control variables that were found to be non-statistically significant were dropped and the model was re-estimated. In this process four control variables were dropped from the model: witness use of drugs or alcohol at the time of the robbery, photo spread administration at the witness's residence, witness did not wear his/her glasses, and involvement in a serial investigation. Removing these variables reduced the problem of cells with zero observations (20% of cells). Substantively, the effects of different photo spread administration methods are unchanged in the full and reduced models. The results presented below (Tables C-5 and C-6) are based on reduced models.

Table C-5. Multivariate multinomial regression results: Effects of photo spread administration method on selection results.

| Photo Spread Comparison                             | Photo Selection Comparison   |   |  |
|---|--|---|--|
|   | No selection or filler selection   | No selection or suspect selection   | Filler selection or suspect selection                                      |
| Blinded sequential compared to blinded simultaneous | A filler selection is more likely when blinded sequential is used            | A suspect selection is more likely when blinded sequential is used            | No difference  |
| Blind simultaneous compared to blinded simultaneous | No difference  | A suspect selection is more likely when blind simultaneous is used (p = .088) | No difference  |
| Blind sequential compared to blinded simultaneous   | A filler selection is more likely when blind sequential used                 | A suspect selection is more likely when blind sequential is used              | No difference  |
| Blinded sequential compared to blind simultaneous   | A filler selection is more likely when blinded sequential is used (p = .096) | No difference   | No difference  |
| Blinded sequential compared to blind sequential     | No difference  | No difference   | No difference  |
| Blind simultaneous compared to blind sequential     | A filler selection is more likely when blind sequential is used              | No difference   | A filler selection is more likely when blind sequential is used (p = .087) |

Model fit:  $-2LL = 344.92$ ,  $\chi^2 = 102.577$ ,  $p = .000$

The results of the reduced multinomial model are presented in Table C-5. The patterns in Tables C-3 and C-5 are roughly equivalent; the patterns in Table C-5 also mirror those in Table 9. The methods, overall, do not perform differently when the likelihood of a suspect identification is compared to the likelihood of a filler selection. The one exception is that filler selections are more likely when a blind sequential method is used rather than a blind simultaneous method is used. The difference between these two methods approaches statistical significance ( $p = .087$ ). The patterns do not clearly suggest one method or set of methods performs better than others by reducing the chances of filler selections. Overall, simultaneous methods are associated with lower chances that a photo will be selected than sequential methods.

Table C-6 summarizes the results of the multivariate model that uses a four category outcome variable: no selection, positive suspect selection, tentative suspect selection, and filler selection. These results are based on a reduced model that includes the same set of control variables as used in the model described in Table C-5 above. The results in Table C-6 are similar to those in Table C-3 and do not suggest that the results are sensitive to inclusion or exclusion of cases that violated the experimental protocols.

Table C-6. Multivariate multinomial regression results: Effects of photo spread administration method on selection results, disaggregated by witness confidence.

|   | Photo Selection Comparison   |  |  |  |  |  |   |
|---|--|--|--|--|--|--|---|
| Photo Spread Comparison                             | No selection or filler selection   | No selection or positive suspect selection   | No selection or tentative suspect selection                                  | No selection or tentative suspect selection                                  | Positive suspect or tentative suspect selection                              | Positive suspect or filler selection   | Tentative suspect or filler selection   |
| Blinded sequential compared to blinded simultaneous | A filler selection is more likely when blinded sequential is used            | No difference  | A tentative suspect selection is more likely when blinded sequential is used | A tentative suspect selection is more likely when blinded sequential is used | A tentative suspect selection is more likely when blinded sequential is used | No difference  | A tentative suspect selection is more likely when blinded sequential is used (p = .084) |
| Blind simultaneous compared to blinded simultaneous | No difference  | A positive suspect selection is more likely when blind simultaneous is used (p = .084) | No difference  | No difference  | No difference  | No difference  | No difference   |
| Blind sequential compared to blinded simultaneous   | A filler selection is more likely when blind sequential is used              | No difference  | A tentative suspect selection is more likely when blind sequential is used   | A tentative suspect selection is more likely when blind sequential is used   | No difference  | No difference  | No difference   |
| Blinded sequential compared to blind simultaneous   | A filler selection is more likely when blinded sequential is used (p = .100) | No difference  | A tentative suspect selection is more likely when blinded sequential is used | A tentative suspect selection is more likely when blinded sequential is used | A tentative suspect selection is more likely when blinded sequential is used | A filler selection is more likely when blinded sequential is used (p = .074) | No difference   |
| Blinded sequential compared to blind sequential     | No difference  | No difference  | No difference  | No difference  | No difference  | No difference  | No difference   |
| Blind simultaneous compared to blind sequential     | A filler selection is more likely when blind sequential is used              | No difference  | A tentative suspect selection is more likely when blind sequential is used   | A tentative suspect selection is more likely when blind sequential is used   | A tentative suspect selection is more likely when blind sequential is used   | A filler selection is more likely when blind sequential is used              | No difference   |

Model fit: -2LL = 453.836,  $\chi^2 = 125.196$ ,  $p = .000$



# Sam Houston<sup>™</sup> STATE UNIVERSITY

MEMBER THE TEXAS STATE UNIVERSITY SYSTEM<sup>™</sup>