| | |
|---|---|
| **Document Title:** | **Reduction of False Convictions through Improved Identification Procedures: Further Refinements for Street Practice and Public Policy** |
| **Author(s):** | **Nancy K. Steblay** |
| **Document No.:** | **249006** |
| **Date Received:** | **July 2015** |
| **Award Number:** | **2007-IJ-CX-0046** |

# Reduction of False Convictions through Improved Identification Procedures: Further Refinements for Street Practice and Public Policy

Nancy K. Steblay

Final Report

Grant #2007-IJ-CX-0046

January 2, 2012

# Reduction of False Convictions through Improved Identification Procedures: Further Refinements for Street Practice and Public Policy

*Abstract*

The project purpose was to enhance the quality and probative value of forensic eyewitness memory evidence acquired through police lineup procedures. Specific objectives were: (1) An updated meta-analytic review of research comparing simultaneous to sequential lineup formats; an evaluation of the *sequential superiority effect* and articulation of factors that moderate this effect; (2) (3) Controlled laboratory testing of the impact on eyewitness accuracy of three individual lineup procedural components: relaxation of the *Yes/No* dichotomous response requirement of the sequential lineup procedure (to allow *I'm not sure* responses); an Appearance Change Instruction to eyewitnesses; and use of multiple identification tasks with the same witness; (3) Collection and analysis of data in collaboration with the Tucson (Arizona) Police Department to compare eyewitness performance on lineup identifications under double-blind simultaneous versus double-blind sequential lineup procedures. This research integrated these three informational components to generate refined recommendations for field practice and public policy. Sequential lineup superiority was established in both laboratory (meta-analysis) and a field test. In addition, the new laboratory data indicate a positive benefit on eyewitness identification accuracy of a not-sure response option for witnesses, only minimal impact of an Appearance Change Instruction, and a significant negative outcome from repeated lineups.

## Reduction of False Convictions through Improved Identification Procedures: Further Refinements for Street Practice and Public Policy

The purpose of this three-part project is to evaluate procedural aspects of police lineups in order to inform best practices. More specifically, the focus of the investigation is the sequential lineup procedure. More than 25 years after the sequential lineup was introduced as a means to reduce false eyewitness identifications (Lindsay & Wells, 1985), and now with over 70 laboratory tests comparing eyewitness identification accuracy under conditions of sequential lineup versus simultaneous lineup format, it is time to ascertain the reliability of the posited *sequential superiority effect* and to bring the sequential lineup to the field for controlled tests.

The project addresses procedural details for best lineup practice in three ways: (1) Existing laboratory data were meta-analyzed to assess eyewitness performance under conditions of sequential and simultaneous lineup formats, to better articulate the key operating components of the sequential procedure and to test the hypothesis of sequential superiority that emerged from an earlier meta-analysis in 2001 (Steblay, Dysart, Fulero, & Lindsay, 2001); (2) Three laboratory experiments examined ancillary elements of current field lineup practice in simultaneous and sequential lineups – an "I'm not sure" witness response option; an Appearance Change Instruction; and multiple identification tasks with the same witness; and (3) A field experiment in Tucson, Arizona, tested double-blind simultaneous versus double-blind sequential lineups using a scientifically-sound experimental design. This report integrates the three components to generate recommendations for field practice and public policy.

### *Background and rationale*

As eyewitness science is brought to practice and policy, collaboration between scientists and law enforcement is beneficial. The project director of this research has worked with County Attorney's offices (Klobuchar, Steblay, & Caligiuri, 2006; Gaertner & Harrington, 2009) and law enforcement in their implementation of double-blind sequential lineups. A primary lesson from these experiences is that purposeful integration of laboratory and field perspectives can facilitate rigorous science *and* effective street practice. Police investigators can provide useful insight about operational challenges and practice preferences, and can help to devise creative remedies for logistical problems that develop under unique jurisdictional circumstances. For example, "functional equivalents" for double-blind testing have been generated for instances in which a true blind lineup administrator is not available (Gaertner & Harrington, 2009). At the same time, laboratory tests can investigate questions that are more satisfactorily answered under controlled conditions in which a specific factor can be evaluated for its unique impact on eyewitness decisions and in which the identity of the offender is known with certainty. The combination of laboratory and field information can offer sound direction for crime investigators when questions arise about lineup protocol.

The discussion around bringing lab-based recommendations to law enforcement street practice has generated many practical questions. And, even with successful implementation of basic lineup reforms, further questions arise about the specific details of recommended procedure and the parameters within which law enforcement departments can tinker with the new protocol to meet their jurisdictional needs -- melding past-preferred methods to the new ones -- without jeopardizing eyewitness accuracy. Some variability in lineup protocol is not a problem, particularly as it may represent natural field practice. The challenge is to identify adjustments that are non-trivial, that have the potential to compromise the accuracy of witness memory, to bring inaccuracies into the investigator's report and misdirect an investigation, or to confound interpretation of the eyewitness's decision. Although field questions may arise from immediate practical and local interests, the results often hold implications for lineup reform nationally and for broader research and theoretical issues.

For example, one aspect of the current project investigates the impact of a police-preferred revision to the prescribed sequential lineup requirement that the witness make a dichotomous "yes" or "no" decision about each photo before moving to the next photo. In the field, witness responses may be more ambiguous ("I'm not sure"), and a concern for some prosecutors is that a lineup administrator's pressure on a witness to respond decidedly to a specific photo (yes or no) may later be argued in court as akin to conducting an improper show-up. A reasonable question is whether a relaxation of the dichotomous response requirement might affect eyewitness decisions.

Simply put, the practical questions involve two broad inquires: *What are the precise procedural details for best practice of double-blind sequential lineups?* And, *how do the outcomes of double-blind sequential procedures compare to the traditional simultaneous outcomes?*

On the latter point, a 2006 study in Hennepin County MN (Klobuchar, et al., 2006) usefully described how double-blind sequential procedures can be brought into field practice, but did not include a comparative control group, thus leaving open the question of how eyewitness performance under the sequential lineup compares to that of the traditional simultaneous lineup under concurrent investigative conditions in the field. Scientists anticipate that law enforcement will see positive results from the sequential lineup, as have been achieved in the laboratory. However, some law enforcement officials seek additional evidence that the recommended changes will significantly aid field investigations. It is desirable to determine whether the pattern of sequential-simultaneous differences in the lab can be replicated under field conditions. The results of an Illinois field study (Mecklenburg 2006) has complicated the policy picture, but also has been roundly criticized for its lack of attention to essential research design requirements (see, e.g., Schacter et al, 2008). Setting aside its more controversial aspects, a clear lesson from the Illinois study is that the next step in the field must be a scientifically-rigorous experimental test of double-blind sequential versus double-blind simultaneous lineup formats.

Scientists and policy makers must respond to inquiries from a solid base of empirical and practical knowledge. Such knowledge comes from converging sources: *existing laboratory evidence, well-designed field studies, and new laboratory explorations.* This project brings all three sources to bear on the challenge of further refining techniques for collection and evaluation of forensic eyewitness evidence.

***Brief review of relevant literature***

DNA exoneration cases have exposed eyewitness error as the predominant factor involved in false convictions (Wells, Malpass, Lindsay, Turtle, and Fulero, 2000). Over a decade ago, this fact propelled joint action among law enforcement, legal professionals, and eyewitness scientists, and resulted in the 1999 publication of "Eyewitness Evidence: A Guide for Law Enforcement" by the National Institute of Justice (Technical Working Group for Eyewitness Accuracy; hereinafter "The Guide"). Psychological science has shown that eyewitness identfications are frequently unreliable and that unintentional police influence can exacerbate witness tendencies toward inaccurate lineup selections. Just as with physical evidence, memory can easily be contaminated with improper handling, and decisions of both the eyewitness and those who evaluate memory evidence can be improperly influenced by contextual non-evidentiary factors (Wells, Memon, & Penrod, 2006). Procedural changes may reduce the likelihood for contaminated memory or poor eyewitness decisions. The NIJ Guide was an important step toward remediation of the problem, providing science-based recommendations for effective collection of eyewitness evidence. The Guide also alerted law enforcement to three developing refinements: sequential lineup presentation format, double-blind lineup administration, and the use of computers for lineup delivery. In the following years, researchers produced a solid body of laboratory evidence that supports the use of double-blind sequential lineups as a means to secure better quality eyewitness evidence (Steblay, Dysart, Lindsay, and Fulero, 2001; Wells, Memon, and Penrod, 2006). And, computer-based delivery of lineups has now entered the field.

**Sequential lineup presentation**. A sequential lineup attempts to remedy a troublesome aspect of the eyewitness decision-making process, *relative*—versus *absolute*—judgment (Wells, 1984). Standard simultaneous-format police lineups present the eyewitness with all lineup members (e.g., six persons) at one time. Under this simultaneous format, eyewitnesses may compare lineup members to determine which most closely resembles the offender in memory, a process of *relative judgment*. If the witness was able to encode and retain a clear memory of the perpetrator and this person is in the lineup (a *culprit-present* lineup), relative judgment may not be a worrisome issue. The concern, however, is whether the eyewitness will recognize the absence of the offender when in fact the suspect in the lineup is not the perpetrator. DNA-exoneration cases – instances in which the actual offender was not in the lineup -- illustrate exactly this problem: witness inability to correctly reject a *culprit-absent* lineup (The Innocence Project). When the culprit is not in the lineup, relative judgment may lead a witness to pick an innocent suspect (Wells, 1993).

To address this problem, the sequential procedure presents the eyewitness with one lineup member at a time and requires the witness to decide whether or not that person is the perpetrator before moving to the next photo. This one-at-a-time procedure is intended to discourage the eyewitness from simply deciding who most resembles the perpetrator, thus forcing a more absolute decision criterion (Lindsay & Wells, 1985).

A 2001 meta-analytic review of laboratory tests demonstrated reliable positive outcomes from use of a sequential procedure (Steblay, et al., 2001). Simultaneous and sequential lineups produce dramatically different choice and accuracy outcomes. Witnesses who view a simultaneous array are more likely to choose a photo from the lineup. When the perpetrator is present, this higher choosing rate boosts correct identifications, perhaps aided by relative judgment. In a culprit-absent simultaneous display, the increased tendency to choose translates into greater risk of identification error. Particularly when the perpetrator is absent from the lineup, the sequential format is preferable, reducing identification errors by 23% and identification of a designated innocent suspect by 2/3 (27% to 9%).

As defined by laboratory research, good lineup practice assumes a single-suspect model (only one suspect in an array of four or more "fillers"—lineup members known to be innocent of the crime) and that the lineup task is the first identification attempt by the witness. It is typical to provide an instruction to the witness that the person who committed the crime may or may not be in the collection of photos to be displayed (an "unbiased" instruction; Steblay, 1997). Specific to the sequential procedure (Lindsay & Wells, 1985), the witness is instructed that the full sequence of photos will be shown even if an identification is made, (and the procedure is conducted in this manner); the witness is unaware of how many photos are in the sequence; photos are presented one at a time with a decision made before going on to the next photo; and the witness is not allowed to go back to earlier photos during the sequence or place photos next to one another.

**Double-Blind lineup administration**. Double-blind experimental procedure, in which neither experimenter nor subject know the subject's treatment condition, is an essential element of sound scientific method used to prevent inadvertent contamination of research results. Interpersonal expectancy effects occur across a broad set of human interactions, necessitating a double-blind method for data accuracy (Harris & Rosenthal, 1985; Rosenthal, 2002; Rosenthal & Rubin, 1978). First noted as essential for lineup administration by Wells in 1988 and confirmed by a broad group of scientists in their lineup recommendations (Wells, Small, Malpass, Fulero, & Brimacombe, 1998) there is wide agreement among eyewitness scientists that the double-blind lineup procedure is crucial in eyewitness procedures (see e.g, Haw & Fisher, 2004; McQuiston-Surrett, Malpass, & Tredoux, 2006; Wells, 2006).

A double-blind lineup protocol (often simply referred to as "blind") helps to manage the inherently suggestive nature of the situation. The detective displaying the photos does not know which photo depicts the suspect or the suspect's position in the lineup. A lineup administrator who does not know the identity of the suspect is unlikely to lead the witness to the suspect through verbal or nonverbal cues, or to selectively perceive and record lineup results. Notification to the witness that the officer does not know which is the suspect affords the additional advantage that the witness is less likely to seek or infer cues from the officer's behavior. An additional benefit of double-blind procedure is achieved if a measure of the witness's confidence is taken at the time of the identification from a blind administrator and before feedback from police or others (Wells and Bradfield, 1998; Douglass and Steblay, 2008).

## Sequential vs. Simultaneous Lineups: Meta-Analysis

The first component of this project involved an update of the 2001 meta-analysis by Steblay, et al., on comparative eyewitness accuracy rates for sequential versus simultaneous lineups. This 2001document has served as the basis for policy decisions, legal reasoning, and analysis (see e.g., Wells, 2006), although some commentators have expressed reservations about what they see to be an underdeveloped empirical base for policy recommendations (McQuiston-Surrett, et al., 2006). Since 2001, laboratory investigators have continued to investigate the sequential lineup, and the number of sequential-simultaneous comparisons has more than doubled, to 72 experimental tests.

Meta-analysis is a widely adopted quantitative technique (Rosenthal, 1991) used to combine the results of multiple studies, each of which has addressed a common question. The analyses search for meaningful patterns across the extant studies. In this case, the analysis will evaluate the hypothesis of a *sequential superiority effect* using existing studies that have tested simultaneous versus sequential lineup format. A primary advantage of meta-analysis is that the cumulative studies can test a phenomenon with substantially increased sample size and examine the impact of factors across studies that may not have been explored within individual studies. The current analyses focus on three issues relevant to the eyewitness research community, to law enforcement, and to policy-makers: a reassessment of the *sequential-superiority effect* in light of new data; a closer examination of the operational specifics of the sequential procedure; and the comparative diagnosticity (cost-benefit ratio) of sequential and simultaneous lineups.

Lineup superiority is defined in this meta-analysis as a higher diagnosticity ratio (more simply called *diagnosticity*; Wells & Lindsay, 1980; Wells & Turtle, 1986). Diagnosticity indicates how much more likely one event is relative to another. In the case of eyewitness identification, this ratio reflects identifications of the culprit to identifications of an innocent suspect. Lineup performance can be evaluated by computing diagnosticity for each of the two lineup formats, simultaneous and sequential. Then, given any two diagnosticity ratios, the higher of the two is stronger evidence for the proposition that the suspect is the culprit. In the legal system, diagnosticity is known as the index of probative value, the tendency to prove or disprove the truth of an allegation. Therefore, greater lineup diagnosticity is a particularly useful index of superiority as it indicates that a witness's decision stemming from a specific lineup format is more probative of guilt.

The results of this meta-analysis, including attention to theoretical and methodological issues, have been peer-reviewed and published for a scientific audience in a much more extensive document (see Steblay, Dysart, & Wells, 2011). The following report focuses more efficiently on the practical aspect of lineup format for investigators: how lineups can be effectively presented and how they compare to traditional practice.

***Method***

Following an electronic search, a review of conference programs, and personal email requests to approximately 50 eyewitness researchers, the resulting papers were screened according to the following criteria: (1) the study provided a statistical test that compared a sequential to a simultaneous lineup format, (2) the statistics required to directly compute a *z*-test and *r* (effect size) were available either within the article or from the author, and (3) the test was for event memory (not a facial recognition paradigm). We excluded tests of within-subject comparisons of the lineup format manipulation, tests in which multiple culprits were positioned in a single lineup, and those in which witnesses were allowed multiple laps through the lineup (although we allowed first-lap data when participants were unaware of a second-lap option). All tests involved single-suspect lineups in which the witness had only one viewing of a lineup, either simultaneous or sequential, for a given culprit.

Forty-nine papers with 72 non-independent tests of sequential versus simultaneous lineup format from 23 different labs were found acceptable based on the above criteria. This set includes 31 tests from studies represented in the 2001 (Steblay et al.) meta-analysis plus 28 new papers that offer 41 new tests. The dataset includes work from 1985 to 2010, representing 13,143 witness-participants, with 55 published (76%) and 17 unpublished tests. The original data were collected in Canada, the United Kingdom, South Africa, Germany, and the United States.

Authors who reported participant samples in generic plural terms such as "undergraduate students" or "community residents" were assumed to have included both male and female participants; under this assumption 100% of the studies included both genders. Sample sizes ranged from 32 to 2529, with a mean of 182.54. (One sample of 2529 is an outlier; with this test removed, sample sizes ranged from 32 to 619, with a mean of 150.) All tests employed photo lineups and 89% of lineups were of size six. The culprit was male in 90% of the stimulus materials. Ninety-eight percent of the tests reported culprit exposure times of less than five minutes.

**Procedure and Statistics.** Two researchers independently reviewed each article, coded moderator variable information, and calculated decision frequencies. Following Rosenthal (1991), the primary statistics computed were *Z* as a test for significant differences between groups and the correlation coefficient *r* as an index of effect size. We elected to follow the Steblay et al. (2001) meta-analysis and the McQuiston-Surrett, et al. (2006) paper in using unweighted values. The mean effect size for a group of hypothesis tests is referred to in subsequent discussion simply as *r*. A useful aspect of *r* in the comparison of eyewitness lineup decisions is that *r* closely approximates the difference in percentages between conditions. For example, an effect size of $r = .05$ to describe the difference in culprit identifications between the simultaneous and sequential lineup conditions will allow one to correctly surmise that the difference between the groups is approximately five percent.

The signed value of *r* and *z* statistics indicates the direction of the obtained results. Positive *r* and *Zma* values denote support of a sequential advantage, that is, eyewitnesses performed better in the sequential lineup condition. Negative *r* and *z* values indicate the opposite: witnesses in the simultaneous lineup condition were more accurate than participants in the sequential lineup condition. See Appendix A for a list of tests. A fail-safe *N* ($N_{fs}$) indicates the number of fugitive tests with null results that would be necessary to overturn a significant outcome. Alpha is set at .05 and all confidence intervals are calculated at 95%.

***Results: All Tests***

The first analysis includes all 72 tests, each comparing the performance of eyewitnesses between simultaneous and sequential lineup formats. Among these studies are tests in which the very best principles of lineup practice were employed, but also some tests in which (usually intentionally) aspects of poor lineup practice were explored. Thus, this set represents the sum of the data, but without consistent attention to best practices. For example, magnified effect sizes are often associated with specific experimental manipulations that are not recommended for practice, such as lineups biased with respect to foil similarity, instruction, or clothing (see, e.g., Blank & Krahe, 2000, Lindsay, Lea, Nosworthy, et al., 1991). This dataset helps to mimic a variety of practices that may exist in real-world investigation, but subsequent subset analyses will better help to define desirable lineup format. This overall analysis is useful in the investigation as a means to better understand the earlier 2001 (Steblay, et al.,) meta-analysis results and to provide a comparative basis for procedural lineup changes.

**Eyewitness decisions: Culprit-present lineups.** Three outcomes are possible for an eyewitness who views a culprit-present lineup: correct identification of the culprit, an incorrect choice of a filler (a known error) or an incorrect rejection of the lineup (no-pick). Culprit identifications from culprit-present lineups are significantly more frequent with the simultaneous lineup, $Zma$ = -9.57, $p < .0001$, $k = 58$, and $r = -.14$, $N_{fs} = 1905$, with a 14% performance advantage (simultaneous lineup: $M = .52$, $CI_{.95}$ [.47, .57]; sequential lineup: $M = .38$, $CI_{.95}$ [.33, .43]; Table 1). Filler pick rate is equal between lineup formats, at 24%; significant choosing rate differences between simultaneous and sequential lineups are therefore represented in the culprit identification rates.

**Eyewitness decisions: Culprit-absent lineups.** An eyewitness who views a culprit-absent lineup will produce one of two outcomes: correct rejection of the lineup (no pick) or a mistaken identification. Correct rejections are 21% higher with the sequential lineup compared to the simultaneous lineup, *Zma* = 16.45, $p < .0001$, $k = 64$, $r = .22$, $N_{fs} = 6339$, (sequential $M = .64$, $CI_{.95}$ [.58, .70]; simultaneous $M = .43$, $CI_{.95}$ [.37, .49]). For this dichotomous accuracy measure, mistaken identifications (and pick rate) in the two conditions are the reciprocal percentages: 36% and 57%.

**Eyewitness decisions: The designated innocent suspect.** Twenty-seven research tests explored eyewitness reaction to an innocent but similar-appearing suspect planted in a culprit-absent lineup. False identification of the designated innocent suspect was significantly more frequent from the simultaneous lineup (M = .25, $CI_{.95}$ [.18, .32]) than from the sequential lineup (M = .13, $CI_{.95}$ [.09, .17]), *Zma* = 7.95, $p < .0001$, $r = .14$, $N_{fs} = 604$.

**Moderator variable: eyewitness age.** (Full moderator analyses are available in the Steblay, et al., 2011 publication.) The 2001 meta-analysis (Steblay et al.) indicated that a sequential lineup was not of benefit when eyewitnesses were children. More recent research (e.g., Memon & Gabbert, 2003a) indicates that older adults have difficulties with lineup decisions, generating considerable mistaken identifications. In the current available tests, older adults and children show a significant advantage of the simultaneous lineup in a culprit-present lineup condition, *r*s = -.25 (five tests) and -.19 (four tests), respectively, and older adults demonstrate a significant sequential lineup advantage in a culprit-absent condition ($r = .27$, from four tests). However, the general lesson from the few available tests is that older witnesses and children make large percentages of errors. When the culprit is absent from the array, the sequential format seems to inhibit choosing somewhat with older adults, but overall these tests show relatively high filler pick rates with both formats: 50% with sequential and 74% with simultaneous lineups. Similarly, in these tests, children chose from both sequential and simultaneous lineups at high rates (84% when the culprit is present, 75% when the culprit is absent), with very high error levels *regardless of lineup format*. In short, children and older adults show significantly different (and profoundly poorer) eyewitness performance compared to the (non-older) adult population regardless of whether they are using the simultaneous or the sequential format. Subsequent analyses exclude tests of older adults and children.

**Summary.** The results for 72 tests of sequential versus simultaneous lineups are remarkably similar to those obtained in the 2001 (Steblay et al.) meta-analysis and provide evidence for reliability of the obtained effects (Table 1). When the culprit is present in the lineup, witnesses in the simultaneous lineup condition make significantly more culprit identifications. When the culprit is not in the lineup, participants in the sequential lineup condition make significantly fewer mistaken identifications. Exceptions to these patterns exist for young children and the elderly, who tend to make substantial errors under either sequential or simultaneous lineup formats.

**The Full Diagnostic Design Dataset ("Full Design")**

Thirteen labs (27 published tests) used, at a minimum, a full 2 X 2 fully-randomized factorial design to explore lineup format effects (sequential/simultaneous lineup format X culprit present/absent lineups) with adult eyewitnesses (see Appendix A). This subset of studies allows cause-and-effect conclusions about the impact on eyewitness decisions of lineup format between comparable culprit-present and absent conditions. Importantly, this *full diagnostic design dataset* also allows protection of subsequent diagnosticity calculations from the influence of uneven research design and diagnosticity ratios with the strong scientific rigor of published work. All tests, published and unpublished, typically are included in meta-analytic calculations in order to work with an increased amount of information, as calculated above. At the same time, an argument can be made for examining only published work as a means to meet *Daubert* criteria (*Daubert v. Merrell Dow Pharmaceuticals, Inc.,* 1993) and the following diagnostic analyses are useful for that purpose. One additional criterion was deemed necessary for inclusion of a test in the full design dataset: that the performance of witnesses must be clearly above chance (see Steblay, Dysart, & Wells, 2011). The pattern of results with the full design dataset (Table 2) is very similar to that obtained with the 72-test dataset (Table 1).

**Pick rates.** The simultaneous lineup produces significantly higher pick rates than does the sequential lineup for both culprit-present and culprit-absent lineups, leading to more culprit identifications when the offender is in the lineup and more mistaken identifications when he is absent. The 24 tests that provide choosing rates for culprit-present lineups indicate an average sequential lineup choosing rate of .61, $CI_{.95}$ [.54, .68] and an average simultaneous lineup choosing rate of .76, $CI_{.95}$ [.72, .80].

**Eyewitness decisions: Culprit-present lineups.** The significant advantage of the simultaneous lineup for correct identifications in the culprit-present lineup condition is trimmed by six percentage points in this full design dataset compared to the full 72-test set, to an average eight percent, with the range of effect sizes from -.32 to +.21 (r = -.08). Confidence intervals overlap somewhat between sequential lineup $M$ = .44, $CI_{.95}$ [.37, .51] and simultaneous lineups $M$ = .52, $CI_{.95}$ [.47, .57].

**Eyewitness decisions: Culprit-absent lineups.** The significant sequential lineup advantage for reduced mistaken identifications within the culprit-absent lineup condition remains virtually the same in this subset of tests (r = .23). Moreover, all but one of 27 effect sizes for culprit-absent lineups are positive, ranging from -.04 to .78—a robust sequential advantage. Mistaken identifications are significantly more frequent from a simultaneous than a sequential lineup; simultaneous $M = .54$, $CI_{.95}$ [.47, .61]; sequential $M = .32$, $CI_{.95}$ [.25, .39]. Mistaken identification of an innocent suspect planted in the lineup is significantly more frequent from a simultaneous than a sequential lineup (simultaneous M = .28, $CI_{.95}$ [.19, .37]; sequential M = .15, $CI_{.95}$ [.07, .23]).

**Diagnosticity.** The sequential lineup produces a diagnosticity ratio of 7.72, the simultaneous lineup a ratio of 5.78. Identification of the suspect from a sequential lineup is 1.34 times more diagnostic compared to an identification from a simultaneous lineup. If the rate of identifying the known-innocent suspect in the culprit-absent condition is used as the ratio denominator, this diagnosticity index yields 1.86 for the simultaneous lineup and 2.94 for the sequential lineup. The sequential lineup is 1.58 times more diagnostic, that is, more probative.

One purpose of this investigation is to highlight system variables, controllable aspects of identification procedures that may affect eyewitness performance (Wells, 1978). Two core system variables—lineup size and use of a cautionary instruction that the culprit "may or may not be in the lineup" (Steblay, 1997)—were implemented almost uniformly across the tests of the full design dataset and thus do not offer meaningful moderator analysis. Additional system variable tests are documented in the Steblay, et al. 2011 document. Two topics demand attention here, however, lineup construction and the "stopping rule."

**Lineup construction method.** Two primary approaches to lineup construction have been employed in the lab. *Match-to-description* is considered by most researchers to be a superior method of constructing a fair lineup (Luus & Wells, 1991), the lineup fillers based on the eyewitness's description of the culprit. A second lineup formation strategy involves a *match-to-culprit* determination, in which lineup fillers for both culprit-present and culprit-absent lineups are based on the culprit's appearance. The comparative results between the two techniques indicate that both *match-to-description* and *match-to-culprit* outcomes align with the common pattern of lineup format differences: the simultaneous lineup produces significantly more culprit identifications from culprit-present lineups and the sequential lineup significantly reduces mistaken identifications from culprit-absent lineups. Notably, a match-to-description method is associated with higher diagnosticity under both lineup formats, but in conjunction with a sequential format shows the highest diagnosticity (10.00) of system variables examined in the meta-analysis.

**Stopping rule.** The witness who views a simultaneous lineup is allowed to compare photos before deciding on any one (or none) of them, and a first inclination to identify a particular photo may be stifled if another lineup member is found to be a closer match to memory. Conversely, in the sequential lineup, it is intended that the witness make a decision one photo at a time, and researchers who employ a "stopping rule" take a first "yes" decision as final and show the witness no additional photos. The stopping rule has become a moot point in the field; to the best of knowledge, no jurisdiction that has implemented the use of sequential lineups has employed a stopping rule. Detectives typically favor showing the entire array, to avoid the appearance of a show-up if the first lineup member is selected or of a truncated lineup if the procedure is terminated at an early identification. Also, if a witness selects an early filler photo, the investigating detective is later reasonably going to ask, "What would have happened when the witness saw the suspect's photo?" Of course witness viewing of the full array allows a witness to change an earlier decision. But, if a witness revokes an earlier filler pick in favor of the suspect in a later lineup position, or conversely, if the witness first picks the suspect and then discredits this initial identification with a change to a filler, this is important and useful information—about the witness's memory, the quality of the lineup fillers, and, along with the witness's comments, the strength of that eyewitness evidence. The latter pattern may alert investigators to the possibility that the suspect is not the perpetrator. Thus, the full lineup display and a record of witness remarks and any reversals in the decision are important to the investigation. A showing of the full array with a complete record also circumvents any problem of suspect position.

For most research teams in this meta-analysis (77%), the lineup photos were continued after an identification (no stopping rule), and, if instructions did not forbid it, a witness could make a second identification or change a decision. A significant sequential advantage in culprit-absent lineups is apparent no matter what the stopping policy. Also, the diagnostic benefit of the sequential lineup (8.43) surpasses the simultaneous lineup (6.34) under either of these strategies. In the culprit-present condition, the difference in culprit identifications between lineup formats and shrinks to 5% for the "continue-to- the-end" studies, a significant difference in effect sizes, $t(24) = 2.05$, $p = .05$. The technique of continuing the lineup to the end of the full series of photos appears to be sound practice.

**Summary**.

The full 72-test dataset from 23 different labs involving 13,143 participant-witnesses yielded overall results that are highly similar to those reported in the 2001 (Steblay et al.) meta-analysis. A sequential lineup significantly reduces mistaken identifications from culprit-absent lineups. A simultaneous lineup produces significantly more culprit identifications when the offender is in the lineup. Also of importance is the fact that there is now a substantial number of published studies (27) from numerous labs (13) that used the full 2 (simultaneous or sequential) X

2 (culprit present or absent) design. These fully randomized studies represent the only database that can reasonably support cause-and-effect claims about the comparative advantage between lineup formats. The full diagnostic design dataset reveals that the difference between sequential and simultaneous lineups in rates of culprit identification is 8% (in favor of the simultaneous lineup) compared to the 15% difference in the 2001 meta-analysis (that did not exclusively use 2 X 2 designs in the analyses), and that the error difference between sequential and simultaneous culprit-absent is 23% (in favor of the sequential lineup). Together, these figures produce a diagnosticity ratio that is higher for the sequential (7.72) than for the simultaneous (5.78) lineup: this is the *sequential-superiority effect*. The full design dataset also reveals higher sequential lineup diagnosticity in the circumstance when a similar-appearing (to the culprit) innocent suspect is in the lineup (sequential 2.94; simultaneous 1.86).

**The Sequential-Superiority Effect.** There are multiple possible sources for the sequential-superiority effect, such as the one-at-a-time display of photos, the witness not knowing how many photos are in the lineup, or the requirement that the witness makes a decision about each photo. A sequential lineup also prohibits the witness from fully determining if some characteristic of one photo (e.g., the color of the background) is unique to the set. It is quite possible that it is the combination of these factors that is important. It was not the purpose of this meta-analysis to tease apart which elements of the sequential procedure contribute to the sequential-superiority effect. All tests in the meta-analysis adhered to the rules of a single-suspect lineup with at least four (usually five) fillers, a restriction to a first identification attempt, and a single viewing of the lineup by a witness. The sequential procedure prohibited the witness from side-by-side comparison of lineup members or return to previous photos, and a yes/no decision for each photo was necessary before moving to the next. Most tests employed the recommended cautionary instruction to the witness that the true perpetrator may or may not be in the lineup. The modal pattern of eyewitness response—increased culprit identifications from the simultaneous lineup and reduced mistaken identifications from the sequential lineup—was generated from a set of tests that adhered to these requirements. In addition, this pattern is associated with a match-to-description lineup construction method and with tests that employed lineup fairness checks for functional/effective size—thus, arguably unbiased lineups. Police prefer to employ the entire lineup rather than stop at an identification, and the meta-analysis suggests that this is a good idea; the difference in correct culprit identifications between sequential and simultaneous lineups is very small (5%) in studies that have allowed witnesses to continue to the end of the lineup. A reliable pattern of eyewitness performance can now be predicted given adherence to these core recommendations.

## Laboratory Experiments

The laboratory component of this project investigated the impact on eyewitness memory accuracy of three aspects of lineup protocol: (1) relaxation of the "Yes/No" dichotomous response demand in the sequential lineup procedure (allowing an "I'm not sure" response) (2) an Appearance Change Instruction to eyewitnesses; and (3) use of multiple identification tasks with the same witness and same suspect. All three studies used similar subject recruitment procedures and laboratory stimuli. All witness-participants took part in Session 1, as described below. Study 3 participants completed a second session, two weeks after the first session.

**Participants**. Participant witnesses ($N$ = 1,415) were above the age of 18 and treated in accordance with American Psychological Association Ethical Standards for Treatment of Human Subjects. Participants were recruited through the existing college research participant pool, signs posted on campus, and announcements in classes, and word of mouth. Participants had a choice of course credit or a $5-$10 payment.

**Stimulus materials**. Session 1 stimulus materials were developed and used in the PD's previous NIJ grant, therefore well-tested (Steblay, 2006). The experiment is delivered to subjects via laptop computer within an Authorware software program. A video of a short (30-second) purse-snatching serves as the stimulus incident. The event was filmed in color with audio, shot from the victim's perspective, and the film provides the witness a close and clear image of the perpetrator's face for 10 seconds. The perpetrator is a male Caucasian perpetrator, age 21. Each 6-person lineup consists of Caucasian males of approximately 18-23 years of age and was constructed through pilot studies to capture moderate physical similarity between lineup members and the perpetrator. A description-matched strategy was employed (Malpass, Tredoux, & McQuiston-Surrett, 2007). Each photo is a full frontal head-and-shoulders view with a gray background. All lineup members including the offender wore their street clothes for the lineup, none matching the clothing worn by the perpetrator at the time of the crime. Each subject views a six-person lineup of sequential or simultaneous format, target-absent or target-present. For Studies 1 and 2, a computer lineup delivery was used and the six fillers rotated in and out of the target-present lineup. The position of fillers and the position of the perpetrator in the target-present array were balanced, with the exception that the offender never appeared in position 1. Changes in this protocol for Study 3 are described below.

**Procedure**. Each subject was run individually with instructions provided through the computer program and by the experimenter during a 15-minute laboratory interaction. Subjects were informed that the study is guided by the National Institute of Justice, therefore to "please pay serious and close attention to the procedures." After a brief introduction, subjects worked through a computer application that requested demographic information, showed the crime scenario, and provided instructions about the upcoming lineup. The experimenter was out of the room during this time. At the conclusion of the crime video, the experimenter re-entered the room and provided additional information, including a verbal cautionary instruction and an appearance-change instruction.

- *In the lineup you're about to see, we are interested in whether you recognize one of the lineup members to be the perpetrator of the crime in the film. However, the lineup may or may not include the perpetrator you saw in the film.*
- *Sometimes the appearance of the offender (e.g., hair style) can change between the event and the lineup. Please keep this in mind.* (This instruction became an independent variable manipulation in Experiment 2.)

The experimenter maintained a blinded status by remaining across the room from the subject, with the laptop computer facing the subject and blocking the experimenter's view of the witness's lineup materials and response sheet. A minor deception was employed to establish the blind status of the experimenter for participants by informing the witness that the specific perpetrator in the witness's film was one of many employed in multiple versions of the film, thus not knowable to the experimenter. In Studies 1 and 2, subjects were thanked and dismissed at the completion of Session 1. In Study 3, the participant was scheduled for a second visit to the lab in two weeks; the purpose of the second visit was not revealed to the witness.

For all three studies, the key dependent variables are witness choice and accuracy. In addition, an index of diagnosticity estimates the probative value of eyewitness identification evidence for each tested lineup procedure: how much more likely it is that the suspect, if identified, is the culprit rather than an innocent person (Wells & Lindsay, 1980; Wells & Olson, 2002); a diagnosticity value of 1.00 indicates that the eyewitness identification has no probative value. When there is no designated innocent suspect, the employed calculation of diagnosticity is a ratio of the percentage of guilty suspect identifications from the culprit-present lineup to the average filler pick rate from the culprit-absent NCD lineup (filler picks/6). Alpha was set at .05, and confidence intervals are reported at the 95% level.

**Laboratory Experiment 1: Relaxation of the *Yes/No* dichotomous response requirement in the sequential lineup procedure (the *I'm not sure* response)** (Experiment 1 has been peer-reviewed and published in the scientific literature; see Steblay & Phillips, in press, for complete analyses.)

As lineup reform continues, it is prudent to determine whether there are additional means to reduce identification errors from eyewitnesses with very weak memories. The sequential lineup push for absolute judgment has been interpreted to mean that a *yes* or *no* dichotomous decision is necessary after each photo. Every one of the 72 tests in the meta-analysis used this dichotomous decision criterion. In the field, however, witnesses are not likely to answer each photo with a clear yes or no. The Hennepin County field study (Klobuchar, et al., 2006) found that a nod, a shrug, or a comment –"*next one*"-- were deemed acceptable as a means for the witness to signal for the next photo. While this practice may seem to undermine absolute judgment, the reality is that any decision to move to the next photo is, by default, a non-identification. Allowance of an *I'm not sure* response may in fact productively reduce guessing. The research question is: Does the inclusion of the *I'm not sure* response option affect witness choosing or accuracy with a sequential (or a simultaneous) procedure?

A forced-choice dichotomous (*yes/no*) decision should not be difficult for a witness who has a good memory of the perpetrator, and minimal impact of an added not-sure response would be expected, at most a slight reduction of positive identifications by some number of witnesses who under a forced-choice protocol would have selected the culprit or an innocent look-alike suspect. Of greater interest is the impact of a not-sure response option on an eyewitness who has a poor memory of the offender—that is, has good reason to be unsure about the lineup decision. Under the prescribed forced-choice sequential protocol, even an uncertain witness must make a firm decision: *yes* or *no*. Neither of these definitive options may effectively capture the true reaction of a witness to a photo, especially a witness with a weakly encoded or diminished memory of the culprit.

Two possible outcomes can be imagined when a not-sure response option is added to this scenario. If the witness would have otherwise opted for a *no* response, the addition of a not-sure option will not change the result; a *not-sure* is the same as *no*. In aggregate, not-sure responses will make no difference to overall witness choosing or accuracy rates. An alternative outcome is that some witnesses who are otherwise willing to pick a lineup member in the absence of true recognition (a relatively low criterion for identification) will switch to the not-sure response, in effect having raised the criterion for identification. This impact would be most apparent in reductions of false alarms; correct identifications may also decrease to the extent that lucky guesses occur with culprit-present lineups.

The prediction for this study is that a not-sure option will increase witness decision accuracy for a sequential lineup. The expectation is that this option will draw some proportion of responses away from *yes* and *no*, preferably from witnesses with weak memories who would otherwise guess.

A secondary purpose of this study is to examine the impact of the not-sure response option on eyewitness performance with simultaneous lineups. Whereas the witness for the sequential lineup must make six decisions, one for each photo, the witness viewing a simultaneous lineup condition makes just one decision for the overall lineup. A not-sure option may inhibit relative judgment during a simultaneous lineup decision in a manner similar to a pre-lineup cautionary instruction ("the culprit may or may not be in the lineup"). However, it can also be posited that the comparative process of relative judgment is so easily and automatically engaged when all six photos are simultaneously displayed, that the salience and impact of a not-sure option will be minimal. The latter scenario is predicted, that the not-sure option used with a simultaneous lineup will have minimal impact on eyewitness performance.

***Method***

The study design is a 2X2X2 factorial, with three independent variables: culprit presence in the lineup (*present/absent*), lineup format (*sequential/simultaneous;* hereinafter *SEQ*, *SIM*) and response options for the witness (*forced-choice/not-sure option;* hereinafter *FC* and *NSO*). Response options were manipulated as the lineup was presented to the witness. For each photo in the SEQ lineup, the *FC* condition included response buttons labeled *yes* and *no*; the *NSO* condition included response buttons labeled *yes*, *no*, and *I'm not sure*. For the SIM lineup, the FC condition included the options to click on a specific photo or to click on a response button labeled *I do not see the person*; the NSO condition included an additional button, *I'm not sure*.

**Participants.** Of 440 participants, 86% self-reported as Caucasian. The remaining 14%, spread across eight experimental conditions, did not allow for a useful analysis of cross-race identification. The results are based on 378 Caucasian witnesses, thereby all same-race identifications, including 55% female, 45% male; with ages ranging from 18-67, *M* age = 21.22 years, *SD* = 5.75.

**Materials.** The fairness of the lineup was assessed: functional size of the lineup was 4.00, effective size was 4.86 (*Tredoux's E',* Tredoux, 1998), [$CI_{95\%:}$ 4.01-6.18]. No a-priori innocent look-alike was included in the lineup. In pilot testing, 55 % of witnesses were able to identify the culprit from a simultaneous array.

**Coding of the not-sure response.** Witnesses who viewed a SIM array examined a single computer screen with six photos, and the witness made one decision for the overall lineup. Witnesses who viewed the SEQ lineup made six decisions, one for each photo. Witness response for both lineup conditions was recorded by the computer as a correct identification, a filler pick, or no-pick. A no-pick response indicated that the witness did not pick any lineup photo from the six (note that the witness may have used the not-sure option for any number of the six sequential photos). A witness's last answer was recorded if multiple picks were made, that is, the computer implemented a last-choice rule.

***Results***

**Manipulation check.** Witness performance is commonly examined in culprit-present versus culprit-absent lineups separately, as reported below. A first comparison, however is between correct identification rate in culprit-present lineups (48.9%) and average misidentification rate in culprit-absent lineups (5.23%), a significant difference, $Z (N = 378) = 9.71$, $p < .001$. Average misidentification rate (31.4/6) was used because there was no a-priori innocent suspect. Rejections of the lineup were significantly more common from a culprit-absent lineup (69%) than from a culprit-present lineup (39%), $Z (N = 378) = 6.00$, $p < .001$. In short, witnesses responded to the perpetrator's photo at better than chance levels, in a pattern expected from the lineup literature (Charman and Wells, 2006). Primary analyses document witness performance at the end of a single sequential lap, with subsequent examination of witness performance in a second sequential lap (See Table 3).

**Witness decisions: Pick rates.** A stepwise logistic regression analysis was used to examine the impact of lineup format and response option and the interaction between these factors upon witness picks of any photo from the lineup. When the culprit was present in the lineup, there was a significant interaction effect, Wald's $X^2$ (1, $N$ = 184) = 21.05, $p < .0001$. The SEQ- NSO condition produced the lowest rate of witness picks (37.0%) and the SIM- FC condition produced the highest rate (83.3%). When the culprit was absent from the lineup, lineup format was a significant predictor of pick rate, with SIM lineups producing more picks than SEQ lineups (46.4% vs. 16.5%), Wald's $X^2$ (1, $N$ =194) = 18.77, $p < .0001$, $r = .31$.

**Witness decisions: Accuracy with culprit-present lineups.** A stepwise logistic regression analysis was used to examine the impact of lineup format and response option (and interaction) upon correct identifications when the culprit appeared in the lineup. There was a significant impact of lineup format, Wald's $X^2$ (1, $N$ = 184) = 15.30, $p <.0001$, $r = .29$. Witnesses viewing a simultaneous lineup were more likely to correctly identify the offender (64.4%) than witnesses who viewed a sequential lineup (35.1%). Response option was a marginally significant predictor ($p = .06$) for culprit identification, with more correct IDs with a forced-choice response option. Filler selections were not significantly affected by lineup format or response option.

**Witness decisions: Accuracy with culprit-absent lineups.** A stepwise logistic regression analysis of the impact of lineup format and response option (and interaction) within culprit-absent lineups revealed a significant impact of lineup format, Wald's $X^2$ (1, $N$ = 194) = 18.77, $p < .0001$, $r = .31$. Witnesses viewing a SEQ lineup more often correctly rejected the lineup (83.5%) than did witnesses viewing a SIM lineup (53.6%). Alternately stated, false alarm rates were significantly higher with a SIM lineup format (46.4%) than with a SEQ lineup format (16.5%).

**Witness decisions: Sequential lineup.** The primary research objective was to explore the impact of a not-sure response option on witnesses who viewed a sequential lineup. Within the SEQ lineup condition, there was a significant difference in pick rate (21%) between FC and NSO conditions when the culprit was present, 58% vs. 37%, $Z$ ($N$ = 97) = 2.10, $p = .02$, $r = .15$. The

NSO condition produced fewer correct identifications, 30% vs. 42%, $Z$ ($N$ = 97) = 1.24, $p$ = .11, $r$ = .11, and fewer false alarms when the culprit was present in the lineup, 7% vs. 16%, $Z$ ($N$ = 97) = 1.50, $p$ = .07, $r$ = .12, and when he was absent, 12% vs. 22%, $Z$ ($N$ = 97) = 1.34, $p$ = .09, $r$ = .12. In combined culprit present and absent lineups, this difference in false alarms became statistically significant, 10% vs. 19%, $Z$ ($N$ = 194) = 2.69, $p$ = .004, $r$ = .12.

**Witness decisions: Sequential laps**. Although not a central component of this research, the computer record allowed assessment of the impact of a second sequential lineup lap. At the completion of the sequential lineup, witnesses were offered the option of a repeat viewing of the same lineup, and 38% of sequential lineup witnesses opted for this second lap. More FC witnesses chose a second lap, 52% vs. 29%, $Z$ ($N$= 194) = 3.29, $p$ = .01, two-tailed, $r$ = .13. The choosing rate increased significantly from first to second-lap (31.4% to 45.9%), McNemar's Test (Binomial), $N$ = 194, $p$ < .0001, pushing up correct identifications in culprit-present lineups (Figure 1) and filler picks in culprit-absent lineups (Figure 2). At the end of two laps, the 12% (non-significant) difference in correct identifications between FC and NSO remained, now at 53% and 41%, respectively. For combined culprit present and absent lineups, false alarms again were significantly higher in the FC condition, 30% vs. 17%, $Z$ ($N$ = 194) = 2.17, $p$ = .02, $r$ = .11.

As might be expected given the aggregate data reported above, analysis of decision changes for individual witnesses from Lap 1 to 2 indicates approximately similar outcomes in FC and NSO conditions. More precisely, witnesses who made no choice in the first lap and then picked from the lineup in the second lap made similar proportions of correct identifications in the FC and NSO conditions , 36% vs. 31%, $Z$ ($N$ = 38) = .33, and similar proportions of false alarms, 41% vs. 27%, $Z$ ($N$ = 37) = .88. In sum, the second lap of the sequential lineup seemingly served the same primary purpose for both FC and NSO witnesses: to allow another chance to make a pick.

However, significantly more of the FC witnesses had already made a lineup pick in Lap 1 and more FC witnesses elected to view a second lap.

**Diagnosticity**. Diagnosticity was compared between groups using an interaction test among proportions (Wells & Olson, 2002). The SEQ lineup produced greater diagnosticity than the SIM lineup (collapsing across FC and NSO conditions), $Z$ ($N$ = 378) = 1.82, $p$ = .03, $r$ = .09. Among the four conditions of the study, the SEQ-NSO was diagnostically superior to the SIM-FC condition. 15.0 vs. 9.2, $Z$ ($N$ = 193) = 1.77, $p$ = .04, $r$ = .10.

**Confidence.** Witness confidence was only slightly higher in the culprit-present lineup condition than in the culprit-absent lineup. Analysis of variance revealed no significant differences in witness confidence across the eight conditions of the study, with witnesses across conditions uniformly settling on a mid-scale response of *somewhat certain* (4 on a 6-point scale).

**Summary.** Three salient aspects of witness performance under a not-sure option were revealed in this study. First, the NSO had a significant impact on witness decisions, in particular reducing the number of witnesses who made a pick from the lineup, but only for sequential lineups. The NSO provided more than a means to simply segregate witness no-choice responses into *no* and *not-sure*; instead, the NSO had a direct impact on witness decisions. Second, the decrease in witness choosing suggests that the NSO raised the identification criterion for some witnesses. As predicted, the sequential lineup coupled with a NSO was revealed to be a more conservative procedure, with a lower false alarm rate and a parallel drop in correct identifications. Interestingly, the purpose of the forced-choice response option is to help secure absolute judgment, yet the relaxation of FC seemingly further raised the bar for making an identification, particularly during a first lap through the sequential lineup. Finally, diagnosticity of a sequential lineup with a not-sure option was significantly higher than the traditional simultaneous forced-choice procedure.

The practical question: Should the not-sure option be employed with a sequential lineup? It is clear that—under the conditions of this study—the NSO reduced lineup picks, including both errors and culprit IDs. Diagnosticity (probative value) favors a single-lap sequential lineup with a not-sure response option. Under these circumstances the likelihood is 15 times greater that a suspect, once picked by the witness, is the offender rather than an innocent suspect. The trade-off between correct identifications and mistaken identifications is a difficult decision for policy-makers. In practice, police may favor a combination of a NSO with an opportunity for a second lap, a practice that reduces diagnosticity. Unlike our laboratory protocol, however, the police have immediate access to the witness for an interview, to fully explore the reasons behind the witness's not-sure answer and/or second lap. The field experiment reported below brings the not-sure response option to the field research protocol.

**Laboratory Experiment 2: The Appearance Change Instruction**

The NIJ Guide recommends that all witnesses be reminded that the perpetrator's appearance may change from the time of crime to the time of lineup photo: "Instruct the witness that individuals depicted in lineup photos may not appear exactly as they did on the date of the incident because features such as head and facial hair are subject to change." (*p*. 32). This Appearance-Change Instruction (ACI) is intuitively appealing as a means to provide context for the witness facing a lineup, but lacks an empirical foundation. ACI effectiveness has been tested minimally in the lab; Charman & Wells (2007) found no benefit to identification accuracy using the ACI with a simultaneous lineup.

Implicit in ACI use is the assumption that correct identifications from culprit-present lineups will be lost when the culprit's appearance in the lineup is not exactly as the culprit looked during the crime event. Also implicit is the assumption that the eyewitness expects a close match between the perpetrator's appearance in the witnessed event and the lineup photo. These reasonable conjectures lead to the prediction that the ACI will urge witnesses to carefully consider features of the lineup members—more so than they otherwise do—to see beyond appearance change, and to recognize the culprit at higher rates than witnesses who are not provided an ACI. In terms of practical outcomes, the ACI should increase correct culprit identifications. This possibility is particularly intriguing as research moves ahead – to establish a means to facilitate memory-based correct identifications as well as to reduce false identifications (Wells, Memon, & Penrod, 2006).

Other scenarios are possible however. Eyewitnesses may, even without an ACI, examine lineup members with awareness that the photos may be outdated or different from recent appearances, thereby making the ACI irrelevant. Alternatively, the ACI may alert the uncertain witness to the difficulty of the identification task and provide a reason for not making a lineup selection, particularly in the case where appearance change is substantial. There is a final possibility: the ACI may only affect witnesses with poor memory, prompting them to lower their threshold for recognition; this would generate more choosing as they allow substantial discrepancy between memory and suspect photo. In practical terms, this scenario would increase choosing rates and severely diminish witness accuracy.

The research question is: *What is the impact of an Appearance-Change Instruction on eyewitness lineup selection rates and identification accuracy in sequential and simultaneous lineups*. It is expected that effects of the ACI will be positive only when the culprit is present in the lineup. No specific prediction is made about the differences expected in simultaneous vs. sequential formats, beyond the expectation from past research that the sequential format will produce a more cautious and accurate eyewitness response.

Two related studies are reported below, with the primary experiment for this research grant reported as Experiment 2b. An earlier study (2a) used the same stimulus event and lineup as the current experiment, with the exception that the culprit's photo involved a hairstyle change (hair combed back off his forehead). Hairstyle is a significant feature for eyewitness identification (Shapiro & Penrod, 1986), and substantial recognition difficulties occurred in Experiment 2a. When a new photo was substituted in additional pilot work (with a similar hairstyle to the crime event portrayal), culprit identifications rose to the level of approximately 50%, similar to that obtained in the current study. The new photo, with culprit hairstyle more similar to the crime event appearance, was used in the primary study reported as 2b below. The juxtaposition of the two studies—one in which the culprit's hairstyle was altered for his photo and one in which the culprit's photo retained a hairstyle similar to the event—provides some information about the functioning of an ACI.

***Method***

Each study is a 2X2X2 factorial study. Three independent variables are manipulated: culprit presence in the lineup (present/absent); lineup presentation format (simultaneous/sequential); and appearance-change instruction (ACI/noACI). Experiment 2a used a culprit photo with a hairstyle change (hair combed back off his forehead; hereinafter, high appearance change); Experiment 2b used a photo with a hairstyle closer to that of the crime event (hereinafter, low appearance change). The two studies were conducted approximately one year apart, therefore the data could not be combined without risks of temporal confounds, and of course the design was not randomized for the factor of appearance change.

**Participants**. 702 participants (504 in Experiment 2a; 198 in Experiment 2b) included a range of ages from 18-65 years, with a mean of 22.51, (*median* = 20.0, *sd* = 7.71). The majority of participants (83.2%) reported themselves to be Caucasian and 59.3% were female.

***Results: Experiment 2a. (High appearance change)***

**Culprit-present lineups**. A logistic regression analysis examined ACI, lineup format, and interactions of these two factors as predictors of culprit identification at the end of the first sequential lineup lap. ACI was a significant predictor of culprit IDs, Wald's $X^2$ (1, $N$ = 230) = 3.95, $p$ = .05 (see Table 4). The ACI produced 11.0% correct identifications; without the ACI, correct identifications were just 4%, a pattern evident across both simultaneous and sequential lineups. There was no significant effect of ACI or lineup format on filler picks within culprit-present lineups. When sequential lineup second laps are included (final decisions for all witnesses), ACI again is a significant predictor of culprit IDs, Wald's $X^2$ (1, 250) = 4.54, $p$ = .03. Lineup format became a marginally significant factor in culprit IDs when second sequential laps are included, Wald's $X^2$ (1, 250) = 2.48, $p$ = .11. In this case, filler picks also were significantly predicted by format, Wald's $X^2$ (1, 250) = 8.97, $p$ = .003. At second sequential laps, both culprit IDs and filler picks were more frequent with sequential than simultaneous lineups.

**Culprit-absent lineups**. A culprit's change in appearance should not affect the culprit-absent lineup, as no photo of the culprit is a part of that lineup. Nevertheless, an ACI may affect the decision process of the witness. A logistic regression analysis was conducted to determine whether lineup format, ACI, and interaction between these factors affected witness's decisions with a culprit-absent lineup. No significant effect for ACI was found (Table 4). At the end of second sequential laps, format was a significant predictor, Wald's $X^2$ (1, 245) = 6.43, $p$ = .01, with greater errors in the sequential than simultaneous lineup condition.

***Results: Experiment 2b (Low appearance change)***

**Culprit-present lineups**. Again, the primary analyses involve simultaneous lineups and sequential lineups at the end of just one sequential viewing. A logistic regression analysis examined ACI, lineup format, and interactions of these two factors as predictors of identification accuracy. The ACI did not significantly affect correct culprit identifications or filler picks, nor did lineup format or the interaction of these factors, Wald's $X^2$ (1, $N$ = 101), $p$s > .40. There were also no significant outcomes at the end of the sequential second lap, $p$s > .10. Post-hoc analysis of ACI impact, separating simultaneous and sequential lineups showed a marginally significant effect of ACI in the pattern of witness responses to simultaneous lineups.

**Culprit-absent lineups**. A logistic regression analysis was conducted to determine whether lineup format, ACI, and interaction between these two factors affected witness's decisions with a culprit-absent lineup. ACI had no impact on witness performance at first or final decision. However, lineup format had a marginally significant impact at final decision, Wald's $X^2$ (1, $N$ = 97) = 2.74, $p$ =.10, with the sequential lineup producing more filler picks, 50% vs. 33%, than the simultaneous lineup.

**Summary.** Experiment 2 determined that an Appearance Change Instruction had no significant impact on eyewitness decisions when the culprit's photo was reasonably similar to his crime event appearance. However, under the condition of a change in culprit's hairstyle, the ACI generated a significant increase in culprit identifications in both simultaneous and sequential lineups. This is the scenario in which the ACI would be expected to yield a positive effect, by correcting for the appearance change. It is clear that the culprit's photo in the high-appearance-change condition was not easily recognized: Without the ACI, most witnesses (74%) did not pick from the lineup, and the correct ID rate did not exceed chance; witnesses who picked were 5.6 times more likely to choose a filler than to choose the suspect. With the ACI, the correct ID rate was four times the no-ACI rate, but still only to a maximum of 12%. Only a third of witnesses who chose from the culprit-present lineup picked the culprit. Unlike the Charman & Wells (2007) investigation of ACI, the current study found no damaging effect of the ACI in either simultaneous or sequential lineups, but a small positive impact under the limited conditions of a very difficult recognition task. The practical lesson from this study is that an ACI may help and is unlikely to hurt identification accuracy.

**Laboratory Experiment 3: Repeated Identification Lineups and Eyewitness Accuracy**

This study tackles the issue of the witness who is asked to attempt a suspect identification a second time. In the field, this situation may play out because of evidentiary requirements. For example, an eyewitness who has successfully identified the suspect from a photo array may be asked to identify the same suspect from a later live lineup, or mug shots or a show-up may be used before a lineup is administered. In England and Wales, if a show-up identification is disputed, a subsequent live lineup must be carried out (Valentine, Davis, Memon, & Roberts, 2011).

Of great concern is the eyewitness who fails to identify the suspect in a first lineup, or does so only tentatively, and later is shown another lineup as a second chance for identification of the same suspect. Larry Fuller was released after 25 years in a Texas prison when he was exonerated by DNA evidence. The victim in the case initially told police that she could not provide a detailed description of her attacker and she was not able to pick him from the first lineup. She later selected him from a second photo lineup; Fuller was the only person present in both lineups (Garrett, 2011, p.59). Garrett (2011) cites at least fourteen DNA-exoneration cases in which the exoneree was the only person repeated in multiple viewings. These instances of second opportunities for identification may have undesirable effects on memory accuracy. If so, policy can and should reflect risk for false identification.

Law enforcement and eyewitness scientists may differently appraise the validity and meaning of a second identification. Common sense may dictate that a repeat identification of the suspect by the witness should boost confidence of both witness and investigator regarding the identification evidence.

The second identification task may involve a better photo, a more up-to-date or vivid appearance of the suspect, thus seemingly a more valid recognition task. In a second identification, law enforcement (and jurors) may see a powerful confirmation of their suspicion about the suspect's guilt. In some cases, a second identification attempt may be prompted only after probable cause has been established (perhaps partially based on the first successful identification) thereby suggesting to the investigator an even greater likelihood of suspect guilt. Yet, it is keenly important to separate the test of witness memory from the very persuasive context of what might be termed a confirmatory environment.

The lineup identification procedures recommended by eyewitness scientists have been developed for a witness's *first* identification attempt (Wells, 2006). This is the test that can best determine whether an eyewitness's recognition memory supports police suspicions about the identity of the perpetrator. Memory confounds inherent to repeated identification practice challenge the fidelity of the witness's memory at a second identification task or beyond. For example, the witness's selection of the suspect at a second lineup may stem from exposure to the same person at the first identification task rather than at the crime scene. A memory failure for the circumstances of the previous encounter (source confusion) results in the witness's sense that the face is familiar although the correct context for that memory has been lost (Brown, Deffenbacher, & Sturgill, 1977). This can manifest as a witness who fails to identify the suspect at the first lineup, but makes a positive identification at the second lineup. A closely related phenomenon is apparent if the witness actively identifies a specific innocent suspect at the first lineup and that new face then supplants that of the perpetrator in memory. This approximates a phenomenon of unconscious transference, whereby a face in close temporal proximity to the offender is confused with the face of the culprit (Loftus, 1976; Ross, Ceci, Dunning, & Toglia, 1994). These two errors of familiarity redirect memory onto a new object, the innocent suspect, at the second lineup.

Subject motivation can drive decision-making, with influences on the outcomes of repeated lineups. If the photo chosen at first exposure is remembered, the witness may feel committed to that selection, presumably to reduce anxiety and to help the investigation; this commitment will be revealed in consistent picks of the same suspect across lineups. Commitment effects have been confirmed in research, such that once a person makes a public decision of identifying a culprit in a lineup, he/she is much more likely to remain with that decision throughout subsequent tests (Dysart, Lindsay, Hammond, & Dupuis, 2001). A commitment effect can be inferred also when choosers continue to choose or when non-choosers continue to not choose; in this case, the commitment is to a decision strategy rather than to a specific lineup member (Goodsell, Neuschatz, & Gronlund, 2009). Finally, the very presence of the same suspect at both identification attempts may suggestively signal to the witness that the police believe this suspect is the perpetrator; the second lineup is the point at which this suggestibility may affect a witness's decision. Admittedly, it is difficult to untangle these factors, and this study will not attempt to do. The interest is in the outcomes of these phenomena, two patterns of eyewitness decisions: consistent identifications across two lineups (suspect/suspect decisions) or decision changes that end with a suspect pick (non-suspect/suspect decisions).

Much of the repeated identification research has focused on identifications that move from a mug-shot or show-up to a later lineup. For example, Goodsell, et al., (2009) explored mugshot commitment effects that result in lineup decision errors. Godfrey and Clark (2010) recently found increased correct and false lineup identification rates from a prior show-up, along with evidence that these effects were due to misplaced familiarity with the show-up suspect and heightened expectations that the show-up suspect would be in the lineup. The most complete work in a relevant area is the meta-analytic review of the *mugshot exposure effect* (Deffenbacher, Bornstein, and Penrod, 2006).

This research reveals that exposure of a suspect to the witness *prior to* the lineup (during examination of mugshots) may result in decreased identification accuracy at the time of a subsequent lineup. In particular, there is an increased probability that the eyewitness will choose a previously-viewed mugshot suspect when the real perpetrator is absent from the lineup.

***Repeated Simultaneous Lineups.*** Experiment 3 explores repeated *lineups*, with an expectation that the nature of the first identification task is important to the phenomenon of repeated identifications. Repeated lineups employ distracters (fillers) at both identification tasks. That is, lineup fillers with physical attributes similar to the offender help protect an innocent suspect by drawing away attention and errant guesses, presumably offering this protection at both lineups. In the same vein of logic, stronger evidence for culpability of a suspect is arguably provided if the witness is able to avoid fillers and select the suspect, dually impressive if achieved at two lineups separated in time. It can be posited that repeated lineups may not present the danger inherent to a mug-shot or show-up pre-lineup identification. Fillers at the first lineup may diminish identifications of an innocent suspect, thus dampening commitment or transference effects that rely on an active pick of the witness at the first identification task. In addition, the salience of an innocent suspect may be reduced by placing him among five fillers of similar features as well as by the passage of time between lineups (the faces of innocent fillers presumably will be less memorable than the perpetrator's) thereby weakening familiarity effects that stem from passive viewing of the suspect at the first identification task.

On the other hand, it must be acknowledged that repeated lineups still possess multiple triggers for memory error—potential transference and commitment effects, and associated problems of memory deterioration over time. Repeated lineups rarely employ the same fillers for both lineups, thus the suspect is still the common member across lineups although somewhat less obvious than with a show-up. These errors may be compounded by witness motivation to simply choose someone from the lineup—a “second chance” orientation that may essentially lower the witness’s criterion for choosing at the second lineup and be exacerbated by diminished memory for the crime. Finally, a simultaneous lineup format allows for relative judgment by the witness, a comparison of lineup members to determine which most closely matches memory (Wells, 1984). It is reasonable to expect that repeated simultaneous lineups (repeated opportunities to engage in relative judgment) may facilitate correct culprit identifications, but also contribute substantially to identification errors. In fact, the situation may be more complex. Wells (2010) points out that the John Jerome White exoneration case contradicts the principle of relative judgment at the second lineup, in that the victim failed to recognize the true perpetrator in the physical lineup (why wouldn’t the true culprit be the best relative match to memory?). Wells sees at least two (not mutually-exclusive) possibilities for this outcome: the victim’s initial mistaken identification of White, presumably the closest lineup member match to her recollection of the crime, changed her memory for the face of the perpetrator (transference), and/or she approached the live lineup with the goal of finding the man she had identified from the photos (commitment). In short, the error of the first lineup repeated itself.

It is this last argument—the potential for errors generated at the first lineup to repeat rather than correct themselves in repeated simultaneous lineups—that forms the hypotheses for this study. The time between lineups in this test—two weeks—is longer than that employed in most studies and likely to result in substantial witness memory deterioration. This retention interval also better approximates the real-life phenomenon of repeated lineups. Despite what is likely to be weakening memories for the culprit over the retention interval, it is predicted that witness choosing rates will be sustained, as witnesses gravitate to the common suspect across simultaneous lineups.

The prediction is that repeated same-suspect simultaneous lineups will increase suspect identifications from first to second lineup. It is likely that guilty-culprit identifications will increase only minimally at second lineup, because existing memory for the culprit and relative judgment will do its work primarily at the first lineup and memory deterioration will limit an increase for true recognition after a two-week interval. More importantly, a significant increase in innocent suspect identifications at a second lineup is predicted. Witnesses who identify an innocent suspect at the first lineup are unlikely to correct this error at the second lineup.

Concern in legal cases and research has focused almost exclusively on repeated identification from traditional simultaneous format lineups. As noted above, the propensity of simultaneous lineups to generate misidentifications in culprit-absent lineups at a first identification task sets the stage for errors to be carried through to a second lineup decision. Extant research on repeated simultaneous lineups supports this prediction (Hinz & Pezdek, 2001; Pezdek, Blandon, and Gitlin, 2005), although these investigations involved a somewhat different scenario from the current study; the intervening lineup was culprit-absent, and the test lineup included both the culprit and innocent suspect.

***Repeated Sequential Lineups.*** The above predictions are made for eyewitness decisions stemming from simultaneous lineups. The sequential lineup appears to be a more conservative procedure, pushing the eyewitness to rely more heavily on absolute judgment (an independent comparison of each lineup member to memory) than on the relative judgment allowed with simultaneous lineups. Witness choosing and misidentification rates are typically lower for sequential than simultaneous format lineups (Steblay, et al., 2011), and it is reasonable to predict that low choosing and error rates at the first lineup will tamp down effects of repeated lineups.

On the other hand, we have seen that repeated *immediate* laps through a sequential lineup significantly increase choosing and identification errors. Research has found that errors ensue particularly from witnesses who did not pick a photo during the first lap, the second viewing became a seeming "second chance" for these witnesses (Steblay, et al., 2011). Thus, there is the potential for a rebound of choosing at the second lineup in this study, albeit two weeks later. Beyond this, the potential for source confusion and common-photo suggestibility are still viable influences with repeated sequential format lineups.

There is an additional aspect of the research design to consider, however. Both simultaneous and sequential lineup response options in this study include the "I'm not sure" response option, explored in Experiment 1. In Experiment 1, the not-sure option appeared to siphon off witnesses with weaker memories and thereby lowered false alarms with sequential lineups, an effect that could benefit sequential lineup outcomes in the current study as memory deteriorates over two weeks. On balance, we anticipate lower initial choosing rates for eyewitnesses who view sequential lineups and only a slight increase in correct and misidentifications from repeated sequential lineups. In sum, we predict a repeated lineup effect for simultaneous but not for sequential lineups.

***Method***

Experiment 3 is a fully crossed randomized 2 *X* 3 factorial design, with two independent variables: lineup format (*sequential/simultaneous;* hereinafter *SEQ, SIM*); and common suspect (*guilty suspect* (GS)/*innocent suspect* (IS)/*no-common-suspect* control group (NCS). Two ancillary simultaneous lineup conditions involved a culprit-absent lineup with the IS only in the second lineup (IS-L2); and a culprit-absent lineup with a new innocent suspect in both lineups (NEW-IS). The rationale for these two conditions is explained in greater detail below. The timeframe for data collection of the two ancillary conditions overlapped with that of the original six. As a precaution, additional data in the SIM-NCS control condition were collected to explore for a possible temporal confound. There was not a significant difference between SIM-NCS data collected at the two (overlapping) timeframes, so the data were collapsed into one SIM-NCS condition. In sum, participants were randomly assigned to eight experimental conditions, and each participant completed two lineups separated by approximately two weeks (*Lineup1, Lineup2*).

***Participants.*** Of 273 male (39%) and female (61%) participants, 77% self-reported as Caucasian. The distribution of non-Caucasian participants across eight conditions did not allow for a meaningful cross-race analysis. Ages ranged from 18-67 years, $M$ = 20.97, $SD$ = 4.99.

***Materials.*** Effective size of the lineups was assessed: *Tredoux's E'* = 4.70, CI [3.99,5.72] for *Lineup1*; and 5.16, CI [4.64,5.81] for *Lineup2* (Tredoux, 1998). Thus, *Lineup2* was a somewhat better lineup in its protection of the suspect. In this study, the lineups were presented in a three-ring binder. The sequential lineup was back-loaded to disguise the number of photos to be seen. The witness's answer sheet included spaces to rate 12 photos (although only six photos were seen). Size of the photos was held constant between simultaneous and sequential displays. Pilot testing indicated that 55 % of witnesses were able to identify the culprit from a simultaneous array; the most frequently chosen photo from a culprit-absent lineup became our innocent suspect (IS).

To counterbalance photo position across subjects, the perpetrator's photo was rotated between positions two and five (switching positions with the filler); all other lineup photos were in fixed position. To mimic this process in the culprit-absent lineups, the innocent suspect (or extra filler in the NCS lineup) alternated between positions two and five. Previous testing found no order effect for these stimulus materials. To counterbalance photo position within subject, the common photo was in a different lineup position from Lineup1 to Lineup2 for a given witness (again switching from position two to five). Pilot testing indicated that choosing rate from the two NCS lineups used at Lineup 1 and Lineup 2 did not differ significantly. Witnesses response options included "I'm not sure."

***Procedure.*** At Session 1, the lineup contained the guilty suspect (GS), an innocent suspect (IS) or a control photo (NCS). After a delay of 14-18 days, the participant returned to the laboratory and was (correctly) informed of our interest in whether his or her memory of the video at this later point in time. The witness was not shown the video in this session, but went directly to the lineup, with instructions as in Session 1. The subjects were shown one of three lineups to match their Session 1 condition: GS, IS, or NCS. All filler photos change from the first to second session.

***Results***

Eyewitness decisions at two points (*Lineup1* and *Lineup2)* are detailed in both Tables 5 and 6. For purposes of brevity, analyses regarding lineup manipulation checks, witness use of the not-sure response option and main effects of lineup format will not be reported here, but are available from the author. In short, witness culprit identification performance was at better than chance levels (a manipulation check), and as expected, SIM lineups generated significantly more choosing than did SEQ lineups at both first and second lineups, producing more culprit IDs with simultaneous lineups and fewer identification errors with sequential lineups. The sequential lineup again produced significantly more not-sure responses than the simultaneous lineup, at both first and second lineups, for both culprit-present and culprit-absent lineups. Diagnosticity was higher for sequential lineups.

**Suspect Picks in Repeated Lineups.** The primary research objective was to explore the impact of repeated lineups. The most direct test of this hypothesis is to compare suspect picks from first to second lineup for all conditions that included a targeted common suspect (the guilty suspect or an innocent suspect). As predicted, overall choosing rates did not change significantly from first to second lineup, in either simultaneous or sequential lineups, McNemar's Test (Binomial), $N = 118$ and 92, respectively, $ps = .21$ and .19. The primary hypothesis was supported: Suspect identifications increased significantly from Lineup1 (31%) to Lineup2 (39%), McNemar's Test (Binomial), $N = 163$, $p = .01$. Separated by suspect type (guilty vs. innocent), the results showed a significant increase in innocent suspect identifications with repeated culprit-absent lineups, 18% to 29%, McNemar's Test, $N = 95$, $p = .007$. Witness picks of the guilty suspect from a culprit-present lineup increased just 3% from Lineup1 to Lineup2, a non-significant difference.

**Simultaneous Lineups: Witness Decisions from First to Second Lineup.** When the culprit was present in repeated *simultaneous* lineups, correct guilty suspect identifications were stable at 65% from first to second lineup, as were filler picks at 11% (Table 5). The identical rates from Lineup1 to Lineup2 hide some inconsistencies in witness responses, however. Eleven percent of witnesses who correctly picked the guilty suspect from Lineup1 did not choose him from Lineup2, and 11% of witnesses who failed to pick the guilty suspect at Lineup1 did choose him from Lineup2. The prediction for (even small) increases in culprit identifications was not supported for repeated culprit-present simultaneous lineups.

When the culprit was absent from the *simultaneous* lineup that contained no designated suspect (NCS), filler picks grew 10% from first to second lineup, McNemar's Test (Binomial), $N = 50$, $p = .20$, suggesting a slight tendency for increased choosing as a function of the second simultaneous lineup, perhaps due to lineup members who were more effectively matched to the culprit or to a witness motivation simply to pick someone at this second chance. Averaged across six lineup members, the risk to any one lineup member was 7% at Lineup1 and 8.67% at Lineup2, a minimal increase. If effective lineup size is used as the denominator, the risk to any one member was 8.94% at Lineup1 and 10.08 at Lineup2. Interestingly, 24% of NCS witnesses misidentified a filler in both lineups (even though all lineup members had changed); an equal number (24%) were consistent in their correct response "I don't see the perpetrator" for both lineups; these patterns may exemplify what Goodsell, et al. (2009) call a commitment-to-strategy.

Most important is risk to the innocent suspect (Table 5). The filler pick rate from a simultaneous culprit-absent lineup with an innocent suspect (IS) increased 13% from first to second lineup, in line with the NCS lineup results (10%), and with study predictions (McNemar's Test, $N = 31$, $p = .19$.). Risk to the innocent suspect increased 16%, McNemar's test, $N = 31$, $p = .09$. These overall changes are not statistically significant, perhaps a function of low power. However, the innocent suspect—who was already at significant risk in Lineup1 compared to any one member of the NCS lineup (36% vs. 7%), $Z (N = 81) = 3.16$, $p = .0008$, $r = .35$—dramatically increased the likelihood that he would be picked when he appeared again in Lineup2, far above any one member of the NCS lineup, $Z = 4.37$, $p < .0001$, $r = .49$. At Lineup 2, his individual risk (52%) matched the combined risk to all members of the NCS lineup (52%). Importantly, of witnesses who chose the innocent suspect at Lineup2, half (47%) had not selected him from Lineup1. The negative impact of a repeated lineup is evident in the diagnosticity index for Lineup2: 1.25 indicates a lineup that is very close to offering no probative information.

**Ancillary Simultaneous Culprit-Absent Conditions.** The increase in risk to the innocent suspect from first to second simultaneous lineups prompts two questions. First, at what rate would witnesses choose the innocent suspect from Lineup2 if he had not appeared in Lineup1? Second, is the risk from double exposure of an innocent suspect limited to a suspect who, due to his physical appearance, is already significantly at risk in Lineup1? To address the first question, witnesses were exposed first to a simultaneous lineup that included neither the perpetrator nor the designated innocent suspect (the NCS lineup). At their return to the lab two weeks later, witnesses were shown a simultaneous lineup that included the innocent suspect. Filler picks increased from first to second lineup, by 20% (McNemar's test, $N = 30$, $p = .09$); filler picks at Lineup1 were somewhat lower than the original NCS Lineup1 condition, and thereby perhaps provided more room for an increase across lineups. More importantly, 30% of witnesses picked the innocent suspect from Lineup2. That is, the innocent suspect was at significantly greater risk in Lineup2, compared to a member of the control group (NCS), $Z (N= 80) = 2.66$, $p = .004$. This proportion of witnesses who chose the innocent suspect when he first appeared in Lineup2 (30%) was slightly less than the proportion in the IS condition who picked him when he first appeared for them in Lineup1 (36%). The innocent suspect's risk was apparent in either lineup, but his predicament as the common suspect across two simultaneous lineups dramatically increased his vulnerability to false identification, compared to being in Lineup1 alone (as established earlier) or in Lineup2 alone, $Z (N = 61) = 1.69$, $p = .05$, $r = .22$.

To address the second question, the innocent suspect was replaced in both lineups with a new innocent suspect (NEW-IS). This lineup member's rate of selection from a culprit-absent lineup in pilot testing was nine percent, not significantly greater than chance. The transition of this former filler to the IS position necessitated adding a new filler to the lineup, thus this lineup is not exactly comparable to those used in other conditions. Nevertheless, the outcome of this condition is informative. To reiterate, the question is whether exposure in two lineups can increase the risk to an otherwise relatively non-salient member of the lineup. The answer: yes. The risk to the NEW-IS more than doubled from first to second lineup, from 12% to 27%, McNemar's Test ($N = 33$), $p = .06$. Although the choosing rate was lower in Lineup2, the innocent suspect became the target of that choosing. Whereas 4 of 13 (31%) filler picks in Lineup 1 were the NEW-IS, nine of ten (90%) of filler picks in Lineup 2 were the NEW-IS, $Z$ ($N$ = 23) = 2.81, $p$ = .002.

**Sequential Lineups: Witness Decisions from First to Second Lineup.** The correct guilty suspect identification rate increased 10% from first to second culprit-present sequential lineup, McNemar's Test (Binomial), $N$ = 31, $p$ = .23, (see Table 6). Here also the overall identification rate hides some inconsistencies of individual witnesses. Fifty-six percent of witnesses who correctly identified the guilty suspect at Lineup1 identified him again at Lineup2; the remaining 44% of those who identified the guilty suspect at Lineup1 did not choose anyone from Lineup2. The increase in correct guilty suspect identifications at Lineup2 was largely due to witnesses who changed their decisions from no-choice (Lineup1) to the guilty suspect (Lineup2). Thirty-three percent of first-lineup non-choosers subsequently picked the culprit.

Culprit-absent *sequential* lineups produced a very low choosing rate (7%) at Lineup1, whether or not the innocent suspect was in the lineup. The choosing rate did not change at Lineup2 for the NCS condition, thus no increase in risk to the innocent suspect. For the IS lineup, the pick rate increased only minimally at Lineup2, up 3% (one additional witness pick). The most common response to a culprit-absent lineup was "not-sure": Seventy-three percent of NCS lineup witnesses and 74% of IS witnesses answered not-sure in one or both lineups. In sum, there is no evidence for a repeated effect in sequential lineups.

**Witness Confidence.** Overall, witnesses were more confident at Lineup1 ($M$ = 4.15, $SD$ = 1.04 [$CI$: 4.03, 4.27]) than Lineup2 ($M$ = 3.86, $SD$ = 1.17 [$CI$: 3.72, 4.00]), $t$ (268) = 3.91, $p$ < .0001. At Lineup2, witnesses who were consistent in their decisions (regardless of the accuracy of the decision) were significantly more confident ($M$ = 4.03, $SD$ = 1.18, [$CI$: 3.82, 4.24]) than were witnesses who changed their decision ($M$ = 3.74, $SD$ = 1.15, [$CI$: 3.56, 3.92]), $t$ (267) = 2.10, $p$ =

.04, two tailed. At Lineup2, witnesses who selected the guilty suspect were significantly more confident ($M$ = 4.34, $SD$ = .97 [$CI$: 4.02, 4.66]) than those who did not ($M$ = 3.65, $SD$ = 1.31, [$CI$: 3.19, 4.11], $t$ (64) = 2.48, $p$ = .02, two-tailed. But at Lineup2, witnesses who selected the innocent suspect from a culprit-absent lineup were also more confident ($M$ = 4.37, $SD$ = .68, [$CI$: 4.06, 4.68] than were those who did not pick him ($M$ = 3.64, $SD$ = 1.12 [$CI$: 3.30, 3.98]), $t$ (59) = 2.60, $p$ = .012, two-tailed. Interestingly, the innocent suspect who was less similar to the culprit (NEW-IS) produced a different pattern of confidence: Witnesses who picked this innocent suspect at Lineup2 were significantly less confident ($M$ = 3.22, $SD$ = .83 [$CI$: 2.68, 3.76] than witnesses who did not pick him ($M$ = 4.29, $SD$ = .86, [$CI$: 3.95, 4.63]), $t$ (31) = 3.21, $p$ = .003, two-tailed.

**Summary.** The hypothesized effect for repeated lineups was supported: Identification of a suspect common to the two lineups increased significantly from Lineup1 to Lineup2. This effect was exclusively a phenomenon of repeated *simultaneous culprit-absent* lineups, producing a significantly increased risk to the innocent suspect at the second lineup. Furthermore, this risk was apparent for both of the tested innocent suspects: one whose appearance put him clearly at risk in the first lineup and one who was a less compelling match to the perpetrator. Witness choosing rates were maintained across the two weeks between first and second lineups, despite a likely deterioration of memory for the culprit, and the picks accumulated on the innocent common suspect at Lineup2. Witnesses who made selections of the guilty suspect or the original innocent suspect at Lineup2 were significantly more confident than witnesses who declined to pick, even when their lineup choice was an error.

Earlier speculation noted the important nature of the first identification task. Unlike a show-up, the first lineup offers a set of fillers to draw witness attention away from the suspect. Nevertheless, it is apparent that lineup fillers did not adequately protect the innocent suspect in

Lineup2, that is, the presence of fillers did not effectively draw errant guesses of witnesses with weak memories away from the innocent suspect. Instead, the repeated simultaneous lineup drew witness misidentifications *to* the innocent suspect.

Repeated culprit-absent simultaneous lineups appear to prompt a mix of witness decision patterns. In some cases, errors made at the first lineup were carried forward to the second lineup; in other cases, errors were generated at the second lineup. Eighty-two percent of witnesses who chose the similar-appearing innocent suspect from the first lineup went on to pick him from again from the second lineup (suspect/suspect decisions), seemingly exhibiting commitment and/or transference effects; this large percentage of witnesses decidedly did not correct their Lineup 1 error. Of witnesses who did not identify the innocent suspect at the first lineup, 35% identified him at the second (non-suspect/suspect decisions). These witnesses may have succumbed to source confusion or the suggestibility of the common photo across lineups. A similar pattern was seen when the innocent suspect was replaced by a less similar (to the culprit) suspect in the culprit-absent lineup (the NEW-IS condition). Seventy-five percent of witnesses in who selected this suspect from the first lineup in this condition picked him again from the second lineup (suspect/suspect decisions); of witnesses who did not pick this suspect from the first lineup, 21% chose him at the second lineup (non-suspect/suspect decisions).

The possible impact of relative judgment is apparent in the comparison of simultaneous and sequential control (NCS) lineups. These culprit-absent lineups offered the witness an opportunity to pick someone from the lineup in the absence of a similar-appearing designated innocent suspect. In the simultaneous lineup condition, 42% of witnesses chose at first lineup, 52% at second lineup, indicating that given the opportunity, substantial numbers of witnesses found some member of the simultaneous lineup to meet the criterion of their memory match. Relative judgment would seem to be in operation. In contrast, very few witnesses in the sequential control lineup condition chose from the lineups, suggesting a limited use of relative judgment and/or a consistently higher criterion for the identification decision.

The research strategy of using the same exact photo in two lineups may have some cost in subsequent application of results to real investigations. However, the sequential lineup results suggest that witnesses did not easily recognize the repeated photo at Lineup 2; In effect, there was no common suspect effect for sequential lineups. Furthermore, the very low choosing and error rate in culprit-absent sequential lineup conditions prompts an intriguing question: Does the lower choosing rate of a sequential lineup also help to prevent the transposing of an innocent suspect's face with the culprit's, thereby lessening the likelihood of retroactive memory distortion? This speculation must be tested in future work. The practical lesson from Experiment 3 is that risk to an innocent suspect is significantly increased when he appears in two identification lineups.

## Tucson Arizona Police Department Field Experiment

The Tucson Police Department (TPD) embarked on a review of their lineup procedures in 2006. Along with the potential for lineup procedural revisions, the TPD is committed to data collection that can inform other jurisdictions across the country. Preliminary discussions between TPD officials and this project's PD produced a research design to compare double-blind sequential and double-blind simultaneous lineups in the field using low-tech tools (e.g., paper lineup photos) but with the required research essentials of proper lineup construction and delivery including true random assignment to experimental conditions. Importantly, the project's double-blind design underscored the importance of documenting filler picks as witness choices, not simply as "failure to ID" and of securing a certainty statement from the witness at the time of the identification and before feedback from a case detective.

As the project developed, the TPD recognized the benefits of joining a wider study on the same topic led by the American Judicature Society (AJS) and its collaborators. In the spring of 2007, TPD made the decision to join the AJS effort. In doing so, Tucson became the second of four cities (Charlotte-Mecklenburg, *N*.C.; Tucson, AZ; San Diego, CA; and Austin, TX) to test sequential versus simultaneous lineup procedures with computer-based administration of the lineups.

The AJS project design followed the Greensboro Protocol, a strategy developed by eyewitness scientists and legal experts in 2006. The Greensboro Protocol outlined a specific and detailed research design for comparison of simultaneous and sequential lineups using laptop computers to implement standardized delivery of the lineups. As noted above, TPD was already positioned to test lineup format using single-suspect six-photo lineups under double-blind administration. AJS provided the resources for laptops, software, and detective training to deliver the lineups in a high-tech version.

The benefits of the software package for the research program are numerous. A photo lineup is constructed by the detective using the police department's existing software program and photo repository; the lineup is then downloaded to the detective's personal laptop computer for showing to the witness at the police station or at a remote location. The laptop computer presents the lineup, thereby ensuring that procedures are administered according to the research protocol (e.g., voice-over and written display of pre-lineup instructions are presented in every instance in a uniform fashion). The computer determines photo order in each lineup with random placement of the suspect in positions 2-6 (never in position1, as requested by the police department) and each witness/lineup is randomly assigned to sequential and simultaneous conditions. This random assignment takes places just after the lineup administrator turns over the computer to the witness for viewing. The program is self-administered (self-paced) by the witness, who can take as long as desired to view the photos. Color photos are equal in size across sequential and simultaneous lineups. All responses of the witness are directly recorded; in the case of the sequential display, response to each of the photos is recorded. The computer permits all the photos to be preserved as part of the electronic record. Once the lineup is complete, the lineup information is uploaded to a server for later collection by the researchers and the case detective.[1]

The Greensboro Protocol requires that the case detective not administer the lineup, therefore each lineup is presented by an administrator who does not know who the suspect is. It should be noted that case detectives have access to the witness and to the computer documentation immediately after the lineup presentation by the blind administrator, specifically for follow-up questions and conversation relevant to the case investigation. The blind administrator also is encouraged to engage with the witness during the lineup procedure to clarify the witness's comments and decision. An ancillary sheet of specific questions for the written record is provided to the blind administrator; this sheet provides a format for the administrator to write down witness comments, particularly statements of identification confidence (e.g., the administrator is cued to question the witness: "In your own words, not using numbers, how confident are you that this is the person you saw?" Detectives were trained by the scientists and liaison PD members.

Case information also becomes a part of the software record, typed into the computer file by the case detective. This information includes crime event factors (e.g., date, type of crime, number of witnesses and perpetrators, viewing conditions, exposure time, weapon use, violence), witness and offender information (e.g., age, race, gender, relationship if any, victim/bystander status, witness's description of the culprit, witness use of drugs), and aspects of the investigation (e.g., time between crime event and lineup, use of any additional ID tasks such as a showup) and lineup procedure (double-blind status, laps through the sequential lineup, lineup photos, etc).

The use of four sites for the AJS study guaranteed that an appropriate number of lineups would be generated for analysis. It was also useful to test communities of different sizes and practices, as a means to identify any problems or obstacles to implementation of computer-delivered double-blind lineups. As it turned out, Tucson experienced some problems in technology and data collection and their experience can inform other jurisdictions about laptop-delivery of double-blind lineups, as discussed below. More importantly, the number of lineups collected in Tucson was far fewer than expected. This report will focus on Tucson; the analyses from all sites are available from the American Judicature Society website.

***Sample***

The Tucson Police Department generated 144 lineups, 45.1% sequential and 54.9% simultaneous. Tucson lineups came from crime investigators in TPD units for robbery, sex crimes, gangs, auto theft, and aggravated assault. Weapons were involved in the crimes for 48% of the reporting lineups. Time between crime event and the lineup ranged from 0 to 679 days (697 is an outlier); 22% of the lineups occurred with one week of the crime and 50% of the lineups within 18 days of the crime. Eyewitnesses were crime victims for 82.6% of reporting lineups (17.4% bystanders).

***Protocol* lineups**. The core set of lineups for analysis (*protocol-consistent* lineups) were double-blind lineups from witnesses who were attempting to identify a stranger and who were seeing a suspect (photo) for the first time since the crime event. In addition, lineups were only included in the *protocol* analyses when the witness made a decision about the lineup (unresolved multiple picks were not included). The computer documentation from each lineup provided the information required to categorize lineups as meeting the *protocol* criteria or not. More specifically:

*Double-blind lineups*: A lineup was considered to be not double-blind if the computer record for any of the following three criteria applied: 1) The lineup administrator answered "yes" to the question "Did you know which image was the suspect?" OR 2) the case detective was also listed as the lineup administrator OR 3) the detective commented in the record that the lineup was not performed double-blind.

*Stranger identifications*: The lineup was considered to not be a case of a stranger identification if the detective answered "yes" to the question of whether the witness knew the perpetrator.

Resolved eyewitness decision: For 10 Tucson lineups (10 of 144 = 6.9%) a witness identified more than one person related to the crime; seven of these were with sequential lineups, three with simultaneous lineups. As with the larger study, if both identifications were fillers, the witness's decision was ruled a filler identification outcome for that lineup. If the witness identified a filler and also the suspect, the researchers—blind to the position of the suspect in the lineup—listened to the audiotape to make a determination as to whether the witness clearly preferred one individual over the other(s); if so, this preference become the final decision for that lineup. If, after reviewing the audiotape, it could not be determined which lineup member was preferred by the witness, the lineup was considered "unresolved" and set aside (2 Tucson lineups).

*First witness viewing of the culprit after the crime*. If the detective documented that the witness had encountered the suspect or the suspect's image after the crime and prior to viewing the lineup (e.g., a picture in the newspaper or on television, or a previous identification attempt such as a live or photo show-up), the lineup was not included in the protocol set.

Of the 144 Tucson lineups, 76 (52.8%) met the criteria for inclusion in the protocol dataset. The remaining 68 lineups did not conform to the protocol for the following specific reasons: Not double-blind, 48; Non-stranger culprit, 10; Witness encountered the suspect or suspect photo between crime and the lineup, 8; and witness's decision unresolved, 2.

***Results* (the *protocol* lineups, *N* = 76)**

Forty-seven of the protocol lineups were simultaneous, 29 were sequential. The problems inherent to small sample size become apparent here, in this lop-sided number of lineups per condition.[2] In small samples sizes, however, just by chance the sample sizes are less likely to be equivalent across conditions.

***Central research question: How does the sequential lineup compare to the simultaneous lineup?***

The central goal of the study is to assess the lineup identification performance of eyewitnesses under conditions of sequential and simultaneous lineups. Witness responses to the lineup can be categorized as a suspect identification, a filler pick, or no pick. It is necessary to note that *suspect identifications* may or may not be accurate decisions. That is, a suspect may or may not be the true culprit. Therefore, we cannot equate suspect identifications with true guilty suspect identifications. Hence, a procedure that produces more or fewer identifications of the suspect might or might not be the best procedure, depending on the proportion of innocent versus guilty suspect picks that each procedure produces. Nevertheless, one of the benefits of true random assignment is that unknown factors—such as the number of guilty or innocent suspects present in the set of lineups—are represented in equivalent numbers in each of the tested conditions. Whatever the true proportion of guilty subjects in these lineups, true random assignment to sequential and simultaneous lineups will distribute this factor evenly between the tested groups. This, of course, is more likely to be true as the sample size increases.

This limited knowledge about the guilty status of a suspect in a lineup does not prevent the conclusion that *filler picks* are clearly mistaken identifications. In this study, every lineup included only one suspect embedded among five fillers, five known-innocent persons. This permitted a test of a central tenet of the sequential procedure, that the sequential procedure reduces filler identifications. Filler identifications serve as a clear proxy index for the relative ability of the two procedures to reduce mistaken identifications. When a witness picks a filler, we know that the eyewitness has made a mistake, that the witness's memory of the culprit is not strong enough to avoid a wrong pick. Also, recognize that once a witness chooses a filler, that witness has lost all credibility as an eyewitness for the case—the witness is spent.

For the 76 Tucson protocol lineups, sequential lineups produced fewer suspect picks, however, this small difference in suspect IDs between the simultaneous and sequential lineups is not a statistically significant difference, Z = .70, $p$ =.58, see Table 7. This means that the difference in suspect identifications from simultaneous versus sequential lineups is within the margins expected by chance using conventional scientific levels of probability; the difference is likely due to chance rather than a true influence of lineup format. On the other hand, rates of filler identifications yield a larger difference, and this 17.4% difference in filler picks is statistically significant, Z = 1.81, $p$ = .04, unlikely due to chance. Simultaneous lineups produced significantly more known errors (filler picks) that did sequential lineups. In short, the simultaneous lineup is significantly more likely to result in a pick from the lineup (the pick rate for the simultaneous lineup is 55%, the pick rate for sequential is 31%); In terms of accuracy, the witness picking from the simultaneous lineup is 2.89 times more likely to make a known error (filler pick). Otherwise stated, when a witness makes a pick from a simultaneous lineup, the pick is just as likely to be a filler as a suspect pick (50% hit rate). From the sequential lineup, a pick is twice as likely to be a suspect as a filler.

***What are the components of an effective sequential lineup?***

The procedural controls of this field experiment provided the basic elements of effective lineup practice, and thereby became the foundation for these results. The protocol lineups were six-person color photo lineups, one suspect per lineup, with different fillers for each suspect. The witnesses did not view a suspect photo (or live suspect) between the time of the crime and the time of the lineup, and each witness viewed the lineup in isolation from other witnesses. The lineups were conducted double-blind. Instructions to the witness included the information that a lineup would never contain more than one suspect and that the person who committed the crime may or may not be included in the lineup. The witness was instructed that s/he did not have to make an identification and that the lineup procedure is important whether or not the witness makes an identification. An appearance change instruction was included for all witnesses. Witnesses were informed that the photos were in no particular order and that they should take as much time as needed. For the sequential lineup, the witness was told that s/he would see photos one at a time, and that the entire series would be shown, even if an identification was made; for the simultaneous lineup the witness was told that the photos would be displayed as a group of six. These are the basic components of a successful sequential lineup.

**Laps**. The study design allowed the witness, upon request, a second viewing ("lap") of the sequential lineup. The lineup administrator could initiate a second lap of the sequential lineup through a password-secure procedure; the lineup was shown again with photos in the same order. The results reported above are for final decisions of the witnesses (including second laps). Only four witnesses in the Tucson study (2.7%) requested a second lap, one witness in the protocol set, three in the non-protocol set. Therefore, the numbers of lap cases are too small for effective analysis in this data set.

**Not sure response option**. The field test included a not-sure response option for the witnesses, which in Tucson was used significantly more frequently by witnesses who viewed a sequential lineup (48.3% used at least one not-sure response) than those with a simultaneous lineup (14.9%). Recall that the not-sure response represents a single decision for the entire simultaneous lineup, but not-sure could be a decision for any of six sequential photos. Again, these Tucson cell sizes are very small, perhaps limiting the usefulness of the results. The analysis of "not sure" responses included an important distinction between two types of non-identification. One type of non-identification is a lineup rejection: the witness never says "yes" to any photo *and* never uses the "not sure" option, "no" is the only answer given. A "not sure" lineup, in contrast, is when the witness never says "yes' to any photo but says "not sure" to at least one photo with the sequential procedure or says "not sure" to the set of photos with the simultaneous procedure. Hence, in this dataset, witnesses using the sequential procedure were less likely to identify a filler, just as effective in identifying the suspect, but also less likely to reject the lineup altogether when they did not make an identification. This could be important for law enforcement and prosecutors as they could consider the suspect as having not been ruled out by the witness.

**The computer software**. The computer-based platform for lineup administration is an important standardization tool of the Greensboro Protocol. However, the question also arises about the practical use of laptop computers for everyday lineup delivery. The Tucson PD provided useful information about their relevant experiences and the numerous IT challenges faced during the field study. As reported by the TPD, the most significant challenge in this study was information technology. An initial time-intensive task for the project was to build a software/server module that allowed the mug photos in the TPD's local database to link with the laptop software used for the study. This process involved collaboration between Mike Garner (*SunGard Public Sector, Inc*) and TPD's IT staff. The interface was built and detectives trained with the computers. However, a few months into data collection, TPD learned that the Arizona Department of Public Safety, after years of requests, had made available driver license photos for lineups. The Motor Vehicle Division (MVD) photos were better quality, up-to-date, and more abundant than Tucson's local photo repository, thus much more useful to investigations. Not surprisingly, detectives frequently defaulted to the MVD photos. However, those photos were not available in the software module TPD had paid to have built for the interface for study laptops. Once again, an interface had to be built, now to bring the MVD photos into the lineup project. Additional time was spent by TPD to secure State approval to employ the MVD photos for the research study.

The economic recession hit the TPD as the rest of the country, and challenges with IT were amplified due to cutbacks, layoffs and centralization (the City moved IT personnel out of the police building and across town). Furthermore, the IT personnel were no longer assigned to Police, but the entire City. At the same time, AJS lost its funding that supported an on-site assistant for the project (a casualty of the Madoff debacle). Many months were lost for data collection. Thankfully, the police department liaison (Captain Stamatopoulos) was gracious to take on additional tasks related to the research.

The software itself was a problem for some detectives, particularly frustrating in the absence of on-site IT staff. There were complaints that the lineup software added work process due to the multiple procedures necessary to capture/load and reload the lineup (create lineup, download from local server to laptop network, then download from network to hard drive on laptop to show lineup, then reconnect to network and upload completed lineup to server).

Beyond IT issues, some obstacles in Tucson were simply strategic, others unavoidable. Available laptops were first distributed to detectives across crime divisions, but in time it became apparent that very few outside the Robbery Division were doing lineups in the field. In an effort to increase the availability of computers to those who needed them, the laptops were redistributed, giving every robbery detective his/her own laptop. The City of Tucson also experienced a reduction in violent crime during this timeframe, a welcome state of affairs for the city, but a challenge for the research project.

After data collection was complete, TPD Captain Stamatopoulos spoke with robbery detectives about the lineup study and the use of laptops. The consensus was frustration with specific laptop issues (e.g., the volume of the speaker was not adequate in some noisy settings) and dislike for the processes involved (e.g., the time-consuming requirement to create duplicate lineups of the same suspect for multiple witness showings) and for the difficulties involved in the interface among photo repository, laptops, and server. The preference was expressed for low-tech paper-style lineups, whether simultaneous or sequential.

The method used in this experiment represents the only field test of eyewitnesses using laptop computers to instruct, administer, and record photographic lineups. The software did a great job for purposes of obtaining pristine research data, but it turned out to be somewhat clumsy and time-intensive for detectives. Because this was the “first generation” of the laptop lineup software, we will look forward to future versions of the software that are more user-friendly for the detectives and the witnesses, and compatible with a variety of platforms rather than only a PC.

The TPD detectives did not like the double-blind requirement. The "fallback" position for circumstances in which a blind administrator could not be located was for the case detective to deliver the lineup, but stay on the opposite side of the computer (i.e., "computer-blinded").

**Summary and next steps**. The Tucson Police Department and its leadership remained dedicated to this project over years of work and despite many setbacks, and the information obtained guided the department in its decision about lineup policy. Captain Stamatopoulos reports that the TPD has decided to require double-blind sequential lineups and will deliver the lineups via laptop. The TPD is working with the same software expert (*SunGard Public Sector, Inc*) to develop a more user-friendly software program.

The lineups from Tucson, even though limited in number, contribute to and align well with the larger dataset from the AJS four-site study.[3] What do these new field data tell police departments and policy makers about lineup procedure?

The data indicate that the sequential lineup produces no significant difference in suspect identifications compare to the simultaneous lineup, when the sequential procedure is operationalized the way it was here. Furthermore, the sequential procedure produced significantly fewer known errors. To the extent that filler identification rates are a reasonable proxy for mistaken identifications of innocent suspects, this means that the sequential procedure will catch fewer innocent suspects in its net. There seems to be no practical reason why real-world lineups cannot be conducted just as easily using the sequential method as they are using the simultaneous method; there are no meaningful differences between the simultaneous and sequential lineup procedures in effort or time on the part of law enforcement. Unfortunately, the Tucson data alone cannot provide a basis for more nuanced analyses of additional aspects of the lineups and witnesses represented in the data.[4]

## Implications for Practice and Future Research

The picture for sequential lineup performance is more sharply in focus as a result of this research project. *Sequential superiority* has been reliably established in the lab, as reported in this project's meta-analysis, and now in the field. The project also has clearly defined operational specifics of successful sequential lineup delivery.

The laboratory studies of this project revealed support for three important principles. First, allowing an "I'm not sure" response option has the potential to strengthen eyewitness decisions with a sequential lineup format. Under the stimulus conditions of this study, the "I'm not sure" response option reduced lineup picks, seemingly providing an off-ramp in decision-making for witnesses who have a weaker memory of the culprit. Second, the Appearance Change Instruction was found to have a small positive effect on culprit identifications, although limited to a scenario in which the culprit's appearance had changed substantially. Importantly, there were no negative effects of the ACI. Further testing of this revision to sequential procedure, using a wider spectrum of stimulus events (Wells & Windschitl, 1999) is necessary to shore up the findings presented here.

A third finding was that the practice of repeated *simultaneous* lineups for the same witness and same suspect significantly increased the risk to an innocent suspect. The probative value of the simultaneous lineup became close to nil at the second lineup. Importantly, of simultaneous lineup witnesses who made a misidentification error (filler or innocent suspect pick) at the initial lineup, the majority (61%) failed to correct that error at the second lineup. Otherwise stated, a Lineup 1 error is 1.56 times more likely to be carried forward as an error than as a correction at Lineup 2. The sequential lineup appears to offer some protection against the dangers of repeated lineups, with even some advantage for increased correct identifications of guilty suspects. This initial favorable evidence should be explored in future research. It must be recognized that the second identification task remains tainted with confounds and thereby does not meet the assumption of the law, that the perpetrator be identified strictly on the eyewitness's uncontaminated memory of the crime scene (Deffenbacher, et al., 2006). The increase in culprit identifications at the second lineup was due to witnesses who had not picked from Lineup 1: one must wonder why the witness could not make an identification at the first lineup when memory was fresh. The specters of transference or suggestibility of common photo loom large. An additional worry with repeated lineup identification procedures of either format is that a witness may be exposed to feedback about his/her decision during the interval between lineups, opening the door for intrusion of post-identification feedback (Wells & Bradfield, 1998; Douglass & Steblay, 2006). Repeated identification tasks are not uncommon in practice (Behrman & Davey, 2001; Garrett, 2011). Yet, to the best of our knowledge, only a few jurisdictions have written policies about repeated identification practices. The State of Wisconsin is an exception, with the following recommendation in its guidelines for law enforcement (Attorney General of the State of Wisconsin, 2005).

> *Avoid multiple identification procedures in which the same witness views the same suspect more than once.* The practice of conducting multiple identification procedures with the same witness and the same suspect should ordinarily be avoided because of the potential for suggestiveness and the potential to contaminate a witness's memory. An eyewitness viewing a second procedure with the same suspect may believe that the suspect's presence in both procedures suggests that authorities believe the suspect is the perpetrator. Or, an eyewitness may become confused and identify the suspect based on recognizing him/her from the prior procedure rather than from remembering the suspect's presence at the crime. In either case, the suggestiveness of the second procedure may

taint the eyewitness. Therefore, eyewitness identification procedures should be approached with the understanding that officers have one opportunity to conduct an eyewitness procedure. Except in unusual cases, conduct only one identification procedure—the most reliable procedure available under the circumstances—in which the same suspect views the same witness once.

Finally, the scientifically rigorous field experiment of sequential versus simultaneous lineups reported here indicates sequential superiority. The sequential lineup produced significantly fewer known errors and no significant reduction of suspect identifications. The real-world conditions of the field test provided meaningful extensions of laboratory conditions—e.g., much longer retention periods between crime and lineup, higher levels of stress and violence, weapon threat, strong victim perspective—yet the data are reasonably consistent with and supportive of sequential lineup effects achieved in the lab. Importantly, the probative value of the sequential lineup procedure is higher than that of the simultaneous lineup procedure.

Implications of this study for police lineup practice are apparent and important. It is hoped that police jurisdictions will consider these data and their implications as lineup procedures are reviewed. A continued dialogue among eyewitness scientists, law enforcement, and legal professionals can prompt additional productive empirical research and effective policy changes. The field data clearly indicate the problems of eyewitness identification. Even with this best-practice protocol, the sequential procedure yielded a 10% rate of filler identifications among those who made a selection from the lineup (the simultaneous was worse, at 28%).

Of greatest importance is the forensic relevance of this project for legal justice. The problem of wrongful conviction must be meaningfully addressed through improved procedures for the collection and evaluation of eyewitness identification evidence. Eyewitness scientists focus on a very important objective – to develop the means to more effectively tap eyewitness memory. This project has provided refinements to secure accurate memory evidence and to build effective policy. For the larger legal and law enforcement communities, the human costs of memory error are particularly close, dramatic, and salient. Erroneous convictions destroy lives, and guilty persons remain unpunished and free to offend again. The practices of most U.S. legal jurisdictions continue to be vulnerable to unacceptable risk of faulty identification evidence (Wells, et al., 2006). Scientists and the legal system must continue to work together to bring genuine improvements for forensic memory evidence into practice and policy.

**References**


American Judicature Society (2011). *A test of the Simultaneous versus Sequential Lineup Method Using Double-Blind Administration with Computers: An An Initial Report of the Eyewitness Identification Field Experiments in Austin, Charlotte, San Diego and Tucson*. Report available at www.ajs.org.

Attorney General of the State of Wisconsin (2005). Model policy and procedure for eyewitness identification. Retrieved from: http://www.doj.state.wi.us/dles/tns/EyewitnessPublic.pdf

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

Blank, H., & Krahe, J. (2000). *The influence of an outstanding similarity between two persons in a lineup on target identification in sequential and simultaneous lineups.* Unpublished manuscript. University of Leipzig, Germany.

Brown, E., Deffenbacher, K., & Sturgill W. (1977). Memory for faces and circumstances of encounter. *Journal of Applied Psychology, 62*, 311-318. Doi: 10.1037//0021-9010.62.3.311

Charman, S., & Wells, G. L. (2006). Applied lineup theory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. *P*. Toglia (Eds.), *Handbook of eyewitness psychology, Volume 2* (pp. 219-254). New Jersey: Lawrence Erlbaum.

Charman, S. D., & Wells, G.L. (2007). Eyewitness lineups: Is the appearance -change instruction a good idea?" *Law and Human Behavior, 31,* 3-22. doi: 10.1007/s10979-006-9006-3.

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

Deffenbacher, K.A., Bornstein, B.H., & Penrod, S.D. (2006). Mugshot exposure effects: retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior*, *30* (3), 287-307.doi:10.1007/s10979-006-9008-1

Douglass, A.B., & Steblay, N. (2006). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology, 20*, 859-869. doi: 10.1002/acp.1237

Dysart, J.E., Lindsay, R.C.L., Hammond, R., & Dupuis,*P*. (2001). Mug shot exposure prior to lineup identification: Interference, transference, and commitment effects. *Journal of Applied Psychology, 86(6),* 1280-1284. doi: 10.1037//0021-9010.86.6.1280

Gaertner , S., & Harrington, J. (2009). Successful eyewitness identification reform: Ramsey County's blind sequential lineup protocol. *The Police Chief, 76*, 130-141.

Garrett, B.L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Cambridge, MA: Harvard University Press.

Godfrey, R.D., & Clark, S.D. (2010). Repeated eyewitness identification procedures: Memory, decision making, and probative value. *Law and Human Behavior, 34*, 241-258. doi: 10.1007/s10979-009-9187-7

Goodsell, C.A., Neuschatz, J.S., & Gronlund, S.D. (2009). Effects of mugshot commitment on Lineup performance in young and older adults. *Applied Cognitive Psychology, 23*, 788-803. doi: 10.1002/acp.1512.

Harris, M.J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin, 97*, 363- 386. doi:10.1037//0033-2909.97.3.363

Haw, R. & Fisher, R.P. (2004).Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology*, 89, 1106-1112. doi:10.1037/0021-9010.89.6.1106

Hinz, T., & Pezdek, K. (2001). The effect of exposure to multiple lineups on face identification accuracy. *Law and Human Behavior, 25*, 185-198. doi:10.1023/A:1005697431830

Klobuchar, A., Steblay, *N*., & Caligiuri, H. (2006). Improving eyewitness identifications: Hennepin County's blind sequential lineup pilot project. *Cardozo Public Law, Policy & Ethics Journal, 4,* 381-413

Lindsay, R. C. L., Lea, J. A., Nosworthy, G. J., Fulford, J. A., Hector, J., LeVan, V., & Seabrook, C. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76,* 796-802. doi:10.1037/0021-9010.76.6.796

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

Loftus, E.F. (1976). Unconscious transference in eyewitness identification. *Law and Psychology Review, 2*, 93-98.

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

Malpass, R.S., Tredoux, C.G., & McQuiston-Surrett, D. (2006). Lineup constructions and lineup fairness. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of Eyewitness Psychology: Memory for People*, (pp. 155-178). Mahwah, N.J., Erlbaum.

McQuiston-Surrett, D. E., Malpass, R. S., & Tredoux, C. G. (2006). Sequential vs. simultaneous lineups: A review of method, data, and theory. *Psychology, Public Policy and Law, 12,* 137-169. doi:10.1037/1076-8971.12.2.137

Mecklenburg, S. (2006a). Report to the Legislature of the State of Illinois: the Illinois pilot program on sequential double-blind identification procedures. Retrieved November 2006 from www.chicagopolice.org.

Pezdek, K., & Blandon-Gitlin, I. (2005). When is an intervening lineup most likely to affect eyewitness identification accuracy? *Legal & Criminological Psychology, 10*, 247-263. doi:10.1348/135532505X49846

Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research*. Newbury Park CA: Sage Publications.

Rosenthal, R. (2002). Covert communication in classrooms, clinics, courtrooms, and cubicles. *American Psychologist*, 57, 834-849. doi:10.1037/0003-066X.57.11.839

Rosenthal, R., & Rubin, D.B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, 3, 377-386. doi:10.1017/S0140525X00075506

Ross, D.F., Ceci, S.J., Dunning, D., & Toglia, M.*P*. (1994). Unconscious transference and lineup identification. *Adult Eyewitness Testimony: Current Trends and Developments* (D.F. Ross, J.D. Read, & M. Toglia, Eds.) Melbourne Australia: Cambridge University Press, 80-100.

Schacter, D. L., Dawes, R., Jacoby, L. L., Kahneman, D., Lempert, R., Roediger, H. L., et al. (2008). Policy forum: Studying eyewitness investigations in the field. *Law and Human Behavior, 32*, 3–5. doi:10.1007/s10979-007-9093-9

Shapiro, P. N. & Penrod, S. D. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*, 139-156. doi:10.1037//0033-2909.100.2.139

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

Steblay, N. (2006). *Double-Blind Sequential Police Lineup Procedures: Toward an Integrated Laboratory & Field Practice Perspective.* Final Report: Grant # 2004-IJ-CX-0044, National Institute of Justice, U.S. Department of Justice.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R.C.L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic review. Law and Human Behavior, 25, 459-473. doi:10.1023/A:1012888715007

Steblay, N.K., Dysart, J. E., & Wells, G.L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law, 17 (1), 99-139*. doi: 10.1037/a0021650

Steblay, N.K., Dietrich, H.L., Ryan, S.L., Raczynski, J.L., & James, K.A. (2011). Sequential lineup laps and eyewitness accuracy, *Law and Human Behavior, 35,* 262-274. doi: 10.1007/s10979-010-9236-2

Steblay, N.K., & Phillips, J. (in press). The not-sure response option in sequential lineup practice. *Applied Cognitive Psychology*. doi: 10.1002/acp.1755

State of Wisconsin, Office of the Attorney General (7-21-06). *Response to Chicago Report on Eyewitness Identification Procedures.*

Technical Working Group for Eyewitness Accuracy (1999) *Eyewitness Evidence: A Guide for Law Enforcement* [booklet]. Washington, DC: U.S. Department of Justice.

The Innocence Project. *The Innocence Project - understand the causes: eyewitness misidentification*. (n.d.). Retrieved from

Tredoux, C.G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior, 22,* 217-237. Retrieved from: http://www.jstor.org/stable/1394144.

Valentine, T., Davis, J.P., Memon, A., & Roberts, A. (2011). Live showups and their influence on a subsequent video line-up. Applied Cognitive Psychology, Published online in Wiley Online Library (wileyonlinelibrary.com). doi: 10.1002/acp.1796

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 12,* 1546-1557. doi:10.1037/0022-3514.36.12.1546

Wells, G. L. (1984). The psychology of lineup identification. *Journal of Applied Social Psychology, 14,* 89-103. doi:10.1037/0022-3514.36.12.1546

Wells, G.L. (1988). *Eyewitness Identification: A System Handbook*. Toronto: Carswell Legal Publications.

Wells, G.L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48(5), 553-571. doi:10.1037//0003-066X.48.5.553

Wells, G. L. (2006). Eyewitness identification: Systemic reforms. *Wisconsin Law Review, 2,* 615-643.

Wells, G.L. (2010). The mistaken identification of John Jerome White. Retrieved from: http://www.psychology.iastate.edu/~glwells/The_Misidentification_of_John_White.pdf

Wells, G.L., & Bradfield, A.L. (1998). “Good, you identified the suspect”: Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83,* 360-376. doi: 10.1037/0021-9010.83.3.360

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

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55,* 581-598. doi:10.1037//0003-066X.55.6.581.

Wells, G.L., Memon, A., & Penrod, S.D. (2006). Eyewitness evidence: Improving its probative value. Psychological Science in the Public Interest, 7 (2), 45-73. doi:10.1111/j.1529-1006.2006.00027.x

Wells, G.L., & Olson, E.A. (2002). Eyewitness identification: Information gain from incriminating and exonerating behaviors. *Journal of Experimental Psychology: Applied, 8*,155-167. doi: 10.1037//1076-898X.8.3.155.

Wells, G. L., & Penrod, S. D. (in press). Eyewitness identification research: Strengths and weaknesses of alternative methods. In B. Rosenfeld, & S. D. Penrod (Eds.), *Research methods in forensic psychology*. New York: Wiley.

Wells, G. L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., & Brimacombe, C.A.E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603-647. doi:10.1023/A:1025750605807

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

Wells, G.L., & Windschitl, P.D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25*, 1115-1125. doi: 10.1177/01461672992512005

**Appendix A**

**Studies in the Meta-Analysis**

* full diagnostic design dataset

Beaudry, J., Mansour, J., Bertrand, M., & Lindsay, R. C. L. (2006, March). *The element of surprise: The impact of participants' knowledge of a subsequent lineup.* Paper presented at the annual meeting of the American Psychology-Law Society, St. Petersburg, Florida.

Blank, H. & Krahe, J. (2000). *The influence of an outstanding similarity between two persons in a lineup on target identification in sequential and simultaneous lineups.* Unpublished manuscript. University of Leipzig, Germany.

*Carlson, C. A., Gronlund, S. D., & Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 14,* 118-128. doi:10.1037/1076-898X.14.2.118

*Clark, S. E., & Davey, S. L. (2005). The target-to-foils shift in simultaneous and sequential lineups. *Law and Human Behavior, 29,* 151-172. doi:10.1007/s10979-005-2418-7

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

Dormer, G. (1983). *Effect of absolute judgmental processes on eyewitness identification accuracy.* Unpublished undergraduate thesis, Queen's University.

Douglass, A. Bradfield, & McQuiston-Surrett, D. E. (2006). Post-identification feedback: Exploring the effects of sequential photospreads and eyewitnesses' awareness of the identification task. *Applied Cognitive Psychology, 20,* 991-1007. doi:10.1002/acp.1253

Dysart, J. E. (1999). *The effects of delay on eyewitness identification accuracy.* Unpublished Masters thesis. Queens University.

Dysart, J. E., & Lindsay, R. C. L. (2001). A pre-identification questioning effect: Serendipitously increasing correct rejections. *Law and Human Behavior, 25,* 155-165. doi:10.1023/A:1005693330921

Ferch, R., & Ebbesen, E. (2003). *A signal detection application to simultaneous and sequential lineups*. Unpublished manuscript, University of California, San Diego.

Gaitens, K. E., Zimmerman, L. A., McQuiston, D. E., & Malpass, R. S. (2002). *Structural bias in eyewitness lineups: When does the suspect "stand out"?* Poster presented at the American Psychology-Law Society biennial conference, Austin, TX.

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

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. doi:10.1037/a0015082.

Hannaford, K. M. (1985). *Relative versus absolute judgments in lineup identifications.* Unpublished undergraduate thesis, Queen's University.

*Kneller, W., Memon, A., & Stevenage, S. (2001). Simultaneous and sequential lineups: Decision processes of accurate and inaccurate eyewitnesses. *Applied Cognitive Psychology, 15,* 659-671. *[This article was in the previous meta-analysis as: Kneller, W., & Memon, A. (2000). Decision processes of accurate and inaccurate eyewitnesses. Paper presented at the American Psychology-Law biennial conference, New Orleans, LA].*

*Levi, A. M. (2006). An analysis of multiple choices in MSL lineups and a comparison with simultaneous and sequential ones. *Psychology, Crime, and Law, 12,* 273-285. doi:10.1080/10683160500238782

Lindsay, R. C. L., & Bellinger, K. (1999). Alternatives to the sequential lineup: The importance of controlling the pictures. *Journal of Applied Psychology, 84,* 315-321. doi:10.1037/0021-9010.84.3.315

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

*Lindsay, R. C. L., Lea, J. A., Nosworthy, G. J., Fulford, J. A., Hector, J., LeVan, V., & Seabrook, C. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76,* 796-802. doi:10.1037/0021-9010.76.6.796

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

*Lindsay, R. C. L., Pozzulo, J., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21,* 391-402. doi:10.1023/A:1024807202926

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

*MacLin, O. H., & Phelen, C. M. (2007). PC_Eyewitness: Evaluating the New Jersey method. *Behavior Research Methods, 39,* 242-247.

*MacLin, O. H., Zimmerman, L. A., & Malpass, R. S. (2005). PC_Eyewitness and the sequential superiority effect: Computer-based lineup administration. *Law and Human Behavior, 29,* 303-321. doi:10.1007/s10979-005-3319-5

Martins, S. (1996). *The effects of changes in the appearance and lineup position of targets on eyewitness identification.* Unpublished undergraduate thesis, Queen's University. Data were also provided by Ferguson (1981), unpublished undergraduate thesis, Queen's University.

*Melara, R. D., Dewitt-Rickards, T. L., & O'Brien, T. *P.* (1989). Enhancing lineup identification accuracy: Two codes are better than one. *Journal of Applied Psychology, 74,* 706- 713. doi:10.1037/0021-9010.74.5.706

Memon, A., & Bartlett, J. (2002). The effects of verbalization on face recognition in young and older adults. *Applied Cognitive Psychology, 16,* 635-650. doi:10.1002/acp.820

*Memon, A., & Gabbert, F. (2003a). Improving the identification accuracy of senior witnesses: Do prelineup questions and sequential testing help? *Journal of Applied Psychology, 88,* 341-347. doi:10.1037/0021-9010.88.2.341.

Memon, A., & Gabbert, F. (2003b). Unravelling the effects of sequential presentation in culprit present lineups. *Applied Cognitive Psychology, 17,* 703-714. doi:10.1002/acp.909

Morgan, III. C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, *P.*, Baranoski, M., Southwick, S.M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry, 27,* 265-279. doi:10.1016/j.ijlp.2004.03.004

Newman, K. (1998). *The effects of a weak memory trace on sequential and simultaneous lineup identifications.* Unpublished undergraduate thesis, Queen's University.

Ogle, C. M., & Reisberg, D. (2006). *A comparison of elimination, sequential, and simultaneous lineup procedures.* Paper presentation at Association for Psychological Science convention, New York, NY.

*Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior, 17,* 11-26. doi:10.1007/BF01044534

Parker, J. F., Tredoux, C., & Nunez, D. (2000). *Lineup measures, lineup procedure, and optimality of encoding.* Paper presented at the American Psychology-Law Society biennial conference, New Orleans, LA.

Phillips, M. R., McAuliff, B. D., Bull Kovera, M., & Cutler, B. (1999). Double-blind photoarray administration as a safeguard against investigator bias. *Journal of Applied Psychology, 84,* 940-951. doi:10.1037/0021-9010.84.6.940

*Pozzulo, J. D., Corey, S. C., Kiraly, M., Girardi, A., Lawandi, A., & Aston, C. (2008). Can a lineup procedure designed for child witnesses work for adults: Comparing simultaneous, sequential, and elimination lineup procedures. *Journal of Applied Social Psychology, 38,* 2195-2209. doi:10.1111/j.1559-1816.2008.00387.*X*

*Pozzulo, J. D., & Marciniak, S. (2006). Comparing identification procedures when the perpetrator has changed appearance. *Psychology, Crime, and Law, 12,* 429-438. doi:10.1080/10683160500050690

Rombough, V. J. (1994). *The effects of lineup presentation and practice on lineup identification accuracy in child eyewitnesses.* Unpublished undergraduate thesis, Queen's University.

*Rose, R. A., Bull, R., & Vrij, A. (2005). Non-biased lineup instructions do matter – a problem for older witnesses. *Psychology, Crime, and Law, 11,* 147-159. doi:10.1080/10683160512331316307

Searcy, J., Bartlett, J. C., & Memon, A. (2000). Influence of post-event narratives, line-up conditions and individual differences on false identification by young and older witnesses. *Legal and Criminological Psychology, 5,* 219-235. doi:10.1348/135532500168100

Shapiro, L. R., & Hiatt, T. A. (2002). *A comparison of photographic line-up procedures.* Paper presented at the American Psychology-Law Society biennial conference, Austin, TX.

Smith, S. M., Lindsay, R. C. L., Pryke, S., & Dysart, J. E. (2001). Postdictors of eyewitness errors: Can false identifications be diagnosed in the cross race situations? *Psychology, Public Policy, and Law, 7,* 153-169. (unpublished sequential lineup data secured from S. Smith). doi:10.1037/1076-8971.7.1.153

Smyth, L. (1994). *Sequential presentation and practice: The jury is still out.* Unpublished undergraduate thesis, Queen's University.

*Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78,* 22-33. doi:10.1037/0021-9010.78.1.22

* Steblay, N.K., Dietrich, H.L., Ryan, S.L., Raczynski, J.L., & James, K.A. (2011). Sequential lineup laps and eyewitness accuracy, *Law and Human Behavior, 35,* 262-274. doi: 10.1007/s10979-010-9236-2

Varrette, T. K. (1994). *The effects of similarity-to-suspect versus match-to-description strategies for selecting lineup foils when perpetrator appearance has been altered*. Unpublished undergraduate thesis, Queen's University.

*Wells, E. C., & Pozzulo, J. D. (2006). Accuracy of eyewitnesses with a two-culprit crime: Testing a new identification procedure. *Psychology, Crime, & Law, 12,* 417-427. doi:10.1080/10683160500050666

*Wilcock, R. A., Bull, R., & Vrij, A. (2005). Aiding the performance of older eyewitnesses: Enhanced non-biased line-up instructions and line-up presentation. *Psychiatry, Psychology and Law, 12,* 129-140. doi:10.1375/pplt.2005.12.1.129

Yarmey, A. D., & Morris, S. (1998). The effects of discussion on eyewitness memory. *Journal of Applied Social Psychology, 28,* 1637-1648. doi:10.1111/j.1559-1816.1998.tb01693.X

Table 1

*Meta-Analysis; 72 tests*

*Lineup Performance: Sequential vs. Simultaneous Lineup Formats*

*2011 Data (Comparative 2001 Meta-analysis Data in Parentheses)* * $p < .05$

| | k | Sequential | Simultaneous | r | *(Zma) |
|---|---|---|---|---|---|
| Eyewitness decisions (Proportions) | | | | | |
| | | 2011 (2001) | 2011 (2001) | | |
| Culprit-present lineup | | | | | |
| Culprit ID | 58 | .38 (.35) | .52 (.50) | -.14 | * |
| Filler | 48 | .24 (.19) | .24 (.24) | | |
| No choice | 48 | .41 (.46) | .27 (.26) | | * |
| Culprit-absent lineup | | | | | |
| Correct rejection | 64 | .64 (.72) | .43 (.49) | .22 | * |
| Filler | 64 | .36 (.28) | .57 (.51) | | * |
| Identification of designated innocent suspect | 27 | .13 (.09) | .25 (.27) | .14 | * |

Table 2

*Meta-Analysis*

*Lineup Performance: Sequential vs. Simultaneous Lineup Formats*

*Full Diagnostic Design Subset of 2x2 Designs, Published, Adult Witnesses*

| | k | Sequential | Simultaneous | r | *(Zma) |
|---|---|---|---|---|---|
| Eyewitness decision (proportions) | | | | | |
| Culprit-present lineup | | | | | |
| Culprit ID | 27 | .44 | .52 | -.08 | * |
| Filler | 24 | .19 | .25 | -.07 | * |
| No choice | 24 | .39 | .24 | | * |
| Culprit-absent lineup | | | | | |
| Correct rejection | 27 | .68 | .46 | .23 | * |
| Filler | 27 | .32 | .54 | | * |
| Diagnosticity ratio | | 7.72 | 5.78 | | |

| | | | | | |
|---|---|---|---|---|---|
| Identification of designated innocent suspect | | | | | |
| | 11 | .15 | .28 | .13 | * |
| Diagnosticity ratio | | 2.94 | 1.86 | | |

Table 3

*Laboratory Experiment 1*

*Eyewitness Decisions (proportions)*

| Response options | | Forced-Choice | | Not Sure Option | |
|---|---|---|---|---|---|
| *Sequential lineup* | | Lap 1 | Lap 2 | Lap 1 | Lap 2 |
| Culprit-present | Correct ID | .42 | .53 | .30 | .41 |
| | Filler pick | .16 | .19 | .07 | .09 |
| | No pick | .42 | .28 | .63 | .50 |
| | *N* | (43) | | (54) | |
| Culprit-absent | Correct rejection | .78 | .58 | .88 | .75 |
| | Filler pick | .22 | .43 | .12 | .25 |
| | *N* | (40) | | (57) | |
| | Diagnosticity | 11.32 | 7.64 | 15.00 | 9.93 |
| *Simultaneous lineup* | | | | | |
| Culprit-present | Correct ID | .69 | | .60 | |
| | Filler pick | .14 | | .13 | |
| | No pick | .17 | | .27 | |
| | *N* | (42) | | (45) | |
| Culprit-absent | Correct rejection | .55 | | .53 | |
| | Filler pick | .45 | | .47 | |
| | *N* | (40) | | (57) | |
| | Diagnosticity | 9.2 | | 7.69 | |

Table 4

*Laboratory Experiment 2a (High appearance change)*

*Eyewitness Decisions (proportions)*

| | Simultaneous | | Sequential (Lap 1) | | Sequential Final | |
|---|---|---|---|---|---|---|
| | ACI | NO ACI | ACI | NO ACI | ACI | NO ACI |
| Culprit-Present | | | | | | |
| Culprit | .10 | .04 | .12 | .03 | .19 | .07 |
| Filler | .20 | .23 | .25 | .21 | .37 | .41 |
| No pick | .70 | .72 | .63 | .76 | .44 | .52 |
| *N* | 70 | 47 | 75 | 58 | 75 | 58 |
| Culprit-Absent | | | | | | |
| Filler | .34 | .38 | .32 | .38 | .53 | .52 |
| No pick | .66 | .62 | .68 | .61 | .47 | .48 |
| *N* | 61 | 55 | 78 | 60 | 78 | 60 |

*Laboratory Experiment 2b (Low appearance change)*

| | Simultaneous | | Sequential (Lap 1) | | Sequential Final | |
|---|---|---|---|---|---|---|
| | ACI | NO ACI | ACI | NO ACI | ACI | NO ACI |
| Culprit-Present | | | | | | |
| Culprit | .57 | .42 | .54 | .54 | .67 | .63 |
| Filler | .04 | .27 | .21 | .13 | .21 | .13 |
| No pick | .41 | .31 | .25 | .33 | .13 | .25 |
| *N* | 27 | 26 | 24 | 24 | 24 | 24 |
| Culprit-Absent | | | | | | |
| Filler | .32 | .35 | .36 | .33 | .56 | .43 |
| No pick | .68 | .65 | .64 | .67 | .44 | .57 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *N* | 25 | 26 | 25 | 21 | 25 | 21 |

Table 5

*Laboratory Experiment 3*

*Eyewitness Decisions from Simultaneous Lineups at Lineup1 and Lineup2*

*Eyewitness decisions (proportions)*

| Lineup Condition | Witness Decision | Lineup 1 | Lineup 2 |
|---|---|---|---|
| Simultaneous Lineups | | | |
| SIM-GS Guilty Suspect (*N* = 37) | Guilty suspect ID | .65 | .65 |
| | Filler | .11 | .11 |
| | No pick | .24 | .24 |
| SIM-NCS Control (*N* = 50) | Correct rejection | .58 | .48 |
| | Filler pick (all) | .42 | .52 |
| SIM-IS Innocent Suspect (*N* = 31) | Correct rejection | .52 | .39 |
| | Filler pick (all) | .48 | .61 |
| | Innocent suspect | .35 | .52 |
| Diagnosticity | | 9.29 (NCS) | 7.47 |
| | | 1.86 (IS) | 1.25 |
| Ancillary Simultaneous Lineups | | | |
| IS-L2 Innocent Suspect in Lineup2 only (*N* = 30) | | | |
| | Correct rejection | .73 | .53 |
| | Filler picks (all) | .27 | .47 |
| | Innocent suspect | ---- | .30 |
| NEW-IS New Innocent Suspect in both lineups (*N* = 33) | | | |

| | | |
|---|---|---|
| Correct rejection | .61 | .70 |
| Filler picks (all) | .39 | .30 |
| Innocent suspect | .12 | .27 |

*Risk to Innocent Suspect in Simultaneous Lineup Conditions*

| | Lineup 1 | Lineup 2 |
|---|---|---|
| Innocent Suspect | .36 | .52 |
| NCS (control) | .07 | .09 |
| IS-L2 | ___ | .30 |
| NEW-IS | .12 | .27 |

GS = Guilty suspect in both lineups

IS = Innocent suspect in both lineups

NCS = No Common Suspect across lineups (control group)

IS-L2 = Innocent Suspect only in Lineup 2 (ancillary condition)

NEW–IS = New Innocent Suspect in both lineups (ancillary condition)

Table 6

*Laboratory Experiment 3*

*Eyewitness Decisions from Sequential Lineups at Lineup1 and Lineup2*

*Proportion selection rates*

| Lineup Condition | Witness Decision | Lineup 1 | Lineup 2 |
|---|---|---|---|
| Sequential lineups | | | |
| SEQ-GS Guilty Suspect (*N* = 31) | Guilty suspect ID | .29 | .39 |
| | Filler | .03 | .03 |
| | No pick | .68 | .58 |
| SEQ-NCS Control (*N* = 31) | Correct rejection | .93 | .93 |
| | Filler pick (all) | .07 | .07 |
| SEQ-IS Innocent Suspect (*N* = 30) | Correct rejection | .94 | .81 |
| | Filler pick (all) | .06 | .19 |
| | Innocent suspect | .06 | .10 |
| Diagnosticity | | 24.16 (NCS) | 32.5 |
| | | 4.83 (IS) | 3.90 |

Table 7

*Field lineup study*

*Eyewitness decisions (percentages)*

*Tucson site (n = 76)*

| | Simultaneous lineup | | Sequential lineup | |
|---|---|---|---|---|
| Suspect ID | 13 | 27.7% | 6 | 20.7% |
| Filler pick | 13 | 27.7% | 3 | 10.3% |
| No pick | 21 | 44.7% | 20 | 69.0% |
| | 47 | | 29 | |

Footnotes

1. A full description of the computer software, the experimental design, and the full program results are available with the AJS summary report at www.ajs.org/.

2. Over the 855 lineups in the full AJS sample, sequential and simultaneous lineups were equally distributed as a result of random assignment by the computer program. Distribution of suspect photos across positions was also equally distributed.

3. Across all four sites of the AJS study (n = 497), pristine lineups produced the same essential results: no significant difference in suspect identification rates between sequential and simultaneous lineups, and significantly fewer filler picks with the sequential lineup.

| | Simultaneous | | Sequential | |
|---|---|---|---|---|
| Suspect ID | 66 | 25.5% | 65 | 27.3% |
| Filler pick | 47 | 18.1% | 29 | 12.2% |
| No pick | 146 | 56.4% | 144 | 60.5% |
| | 259 | | 238 | |

4. The AJS database, of which Tucson is a part, does have a wealth of information from 855 lineups. Future analyses will attempt to explore aspects of the non-protocol lineups (in particular, the non-blind and familiar-culprit subsets), estimator variables related to the crime event (weapon presence, cross-race crimes, delay between crime and lineups, etc.), and aspects of witness decision variables (time spent to decide). Of particular interest is the confidence expressed by the witness in his/her own words (on audiotape files) and its relationship to the witness's decision. Results of these analyses will come out in later reports and in refereed scientific journal articles.