Effect of Retention Interval on Showup and Lineup Performance

Stacy A. Wetmore[1], Jeffrey S. Neuschatz[2], Scott D.Gronlund[1], Alex Wooten[2], Charles A. Goodsell[3], & Curt A. Carlson[4]

[1]University of Oklahoma

[2]The University of Alabama in Huntsville

[3]Canisius College

[4]Texas A&M University – Commerce

Please address editorial correspondence to:

Stacy A. Wetmore
University of Oklahoma
Department of Psychology
455 W. Lindsey Street
Dale Hall Tower, Room 705
Norman, OK 73019-2007
Email: stacy.wetmore@ou.edu

Author Notes

This work was supported by the National Science Foundation (NSF) grant SES–0719997 to Scott Gronlund and NSF grants SES-1060913 to CAG, SES-1060921 to JSN, and SES-1060902 to SDG. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the NSF.

Abstract

Showups – when a single suspect is presented to an eyewitness – are thought to be a more suggestive procedure than traditional lineups by the U.S. Supreme Court and social science researchers. The present experiment examined the impact of retention interval on showup identifications, because immediate showups might be no worse than, and perhaps even better than, a lineup conducted after a delay. Participants (N = 1486) viewed a mock-crime video and then were presented with a showup or a simultaneous lineup, either immediately or a 48-hour delay. ROC (Receiver Operating Characteristic) analyses revealed that a showup never resulted in better identification accuracy than a lineup. We conclude with a discussion of whether showups should ever be used.

Keywords: Showups, Simultaneous Lineups, Receiver Operating Characteristic (ROC) analysis

Effect of Retention Interval on Showup and Lineup Performance

The Supreme Court (*Stovall v. Denno*, 1967; *United States v. Wade*, 1967), state courts (*Bradley v. State*, 1980; *Commonwealth v. Carter*, 1979), and social science researchers (Gronlund et al., 2012; Steblay, Dysart, Fulero, & Lindsay, 2003) have stated that showup identifications are more suggestive than lineup identifications. Although these entities have deemed that showups may be a less accurate identification procedure than lineups, they are still a very common practice among police departments (e.g., Flowe, Ebbesen, Burke, & Chivabunditt, 2001; Garrett, 2011). Therefore, it is important to thoroughly examine this form of identification and determine under what circumstances it can assist law enforcement. In particular, police are often faced with the possibility of presenting a showup to an eyewitness shortly after the crime, or constructing a lineup after a delay. With this application in mind, we had three objectives: (a) to compare identification accuracy between showups and lineups within a single study, (b) to compare these procedures both immediately after a mock crime as well as over a retention interval (48 hours), and (c) to make these comparisons using a robust method (i.e., ROC curves) only recently applied to eyewitness identification procedures.

Identification Procedures

Throughout this paper we utilize several terms imperative to understanding the relationship between showup and lineup identification procedures. A simultaneous lineup presents an array of six or more individuals. *Perpetrator Present* (PP) lineups contain the actual culprit; *Perpetrator Absent* (PA) lineups instead contain an individual who resembles the perpetrator – a designated innocent suspect (i.e., an innocent person who is falsely thought to have committed the crime). Correct identifications occur when a witness correctly chooses the perpetrator from a PP procedure (the guilty suspect); false identifications occur when an innocent

suspect is mistakenly identified from a PA procedure. Identification of any filler (i.e., non-suspect) from a lineup is not deemed a dangerous error because these individuals are known to be innocent before being placed into a lineup.

A showup, in contrast, is an identification procedure in which a single person is presented to the witness, either live or in a photograph (Dysart & Lindsay, 2007; Valentine, Davis, Memon, & Roberts, 2012). Typically, these one-on-one confrontations occur in the field in close spatial and temporal proximity to the crime (Behrman & Davey, 2001; Valentine et al., 2012). The rationale behind showups is that they provide a quick means for detaining the guilty or exonerating the innocent. Just as in lineups, showups can be PP and PA, and again we classify the identifications made from showups as either correct or false identifications, respectively. Unlike lineups, however, there is no opportunity for filler identifications because no other individuals are presented with the guilty or innocent suspect. The lack of known innocents (i.e., fillers) is thought to be one of the problems with this method of identification.

Showups versus Lineups

Given that showups are one of the most common identification techniques (Behrman & Davey, 2001; Garrett, 2011), it is surprising how little research has examined showups relative to the immense literature on lineups. It is possible that the lack of research on showups stems from the fact that legal scholars have accepted that one-person identifications are biased; consequently, researchers do not need to attempt to empirically assess its utility (Gonzales, Ellsworth, & Pembroke, 1993). However, the limited research comparing lineups and showups has produced some conflicting results (see Goodsell, Wetmore, Neuschatz, & Gronlund, 2013). Although the majority of these studies have found that showups are less diagnostic of suspect guilt (Gronlund et al., 2012; Lindsay, Pozzulo, Craig, Lee & Corber, 1997; Steblay et al., 2003;

Wagenaar & Veefkind, 1992), others have found a showup advantage (Flowe & Ebbesen, 2007; Dekle, 2006; Beal, Schmitt, & Dekle, 1995; Dekle, 1997; Dekle, Beal, Elliott, & Huneycutt, 1996), and still others found no significant difference (Valentine et al., 2012). Thus, the research remains equivocal as to which identification procedure is superior, or under which conditions showups might be favored.

Clark and Godfrey (2009) found that correct identification rates did not significantly differ between showups and lineups. The authors suggested, however, that focusing on correct identifications and correct rejections put lineups at a disadvantage. For example, a witness can choose a filler from a lineup but not from a showup. Consequently, correct rejections in PA lineups are reduced because each filler identification is no longer counted as correct rejection, but as an identification. To alleviate this issue, Clark and Godfrey (2009) calculated the Innocence Risk, which is the probability that a suspect is innocent, given that a suspect (guilty or innocent) was identified: $innocence\ risk = pa/(pa + pp)$, where *pa* stands for the probability of choosing the innocent suspect from the PA lineup and *pp* stands for the probability of choosing the guilty suspect from the PP lineup. In their innocence risk analysis, Clark and Godfrey found that, even though showups and lineups did not differ significantly with respect to correct and false identifications, showups still put innocent suspects at a greater risk of being identified falsely.

Retention interval, however, was a moderating variable in this analysis. Clark and Godfrey (2009) reported a lower Innocence Risk for showups that were conducted immediately (.20) than lineups conducted the next day (.30). Hence, it is possible that showups may provide an advantage over lineups when they can be conducted quickly (i.e., when eyewitness memory is still fresh). It is important to note that, according to the courts, showups are only justified when

the suspect is found near the crime scene shortly after the crime occurred (see Manson *v.* Braithwaite, 1977).

Only three studies have compared showups and lineups over a retention interval (Dekle, 1997; Valentine et al., 2012; Yarmey, Yarmey, & Yarmey, 1996). Yarmey et al. presented a live mock crime followed by a showup or lineup either immediately or after a retention interval (30 min, 2 hours, or 24 hours later). In general, eyewitnesses were less accurate as the retention interval increased. More specifically, the rate of false identifications increased in showups as the retention interval increased; however, showups conducted immediately provided more correct identifications relative to lineups performed at a delay. However, one potential limitation of the study was the use of two sisters as the perpetrator and innocent suspect. It is doubtful that a criminal and an innocent suspect would resemble each other that closely; therefore, these results should be interpreted with caution.

Dekle (1997) also found that identification accuracy generally declined with retention interval for both showups and lineups, but signal detection analysis revealed a *d'* (i.e., accuracy) advantage for showups immediately after the crime, after 2-3 days, and after one week. The greatest advantage was for conducting immediate showups versus delayed lineups. Valentine et al. (2012, Experiment 3) also found an advantage for immediate showups relative to delayed lineups. Moreover, they found that participants made fewer false identifications in immediate showups than from lineups. However, there was no statistical difference in false identifications of the innocent suspect from showups conducted immediately compared to lineups conducted after a retention interval.

In sum, the data are not definitive regarding a typical situation encountered by police: whether to present a showup shortly after a crime or a lineup after a delay. If in fact an

immediate showup has greater probative value than a delayed lineup, then it warrants additional and more ecologically-valid investigation. The current study was designed to examine this issue by comparing both identification procedures immediately after the crime and over a retention interval.

Lineup Composition

Retention interval is not the only factor to consider when evaluating lineups and showups. Lineup composition is another factor that could differentially affect eyewitness performance as a function of lineup procedure (e.g., Carlson, Gronlund, & Clark, 2008; Gronlund et al., 2009). Lineup composition refers to the degree to which the fillers in the lineup match the perpetrator. If the fillers are poor matches to the perpetrator, then the lineup is biased because there is a greater chance that the guilty or innocent suspect will be chosen even if the eyewitness has poor memory of the perpetrator. However, if the fillers are reasonable matches to the perpetrator, the lineup is fairer because an eyewitness with poor memory is equally likely to choose a filler or the perpetrator (Luus & Wells, 1991). Although Wells and Quinlivan (2009) suggested that a showup may put an innocent suspect at greater risk than a fair lineup, a showup may still be better than placing an innocent suspect in a biased lineup. To test this hypothesis, we utilized two levels of lineup fairness, biased and fair. The fair lineup contained good matches to the perpetrator; the biased lineups contained poor matches to the perpetrator. [1]

Receiver Operator Characteristic (ROC) Analysis

Recently, it has been argued that ratio-based probative value measures (e.g., the aforementioned Innocence Risk or diagnosticity estimates like correct identifications/false

[1] A third type of lineup was constructed to simulate the way police department procedure. These data were excluded for ease of exposition. Although the data from this lineup behaved similarly to the fair lineup (correct and false identifications rates and pattern of results), collapsing it with the fair lineup would be improper due to different methods of construction.

identifications) are inappropriate for determining which identification procedure is superior (Mickes, Flowe, & Wixted, 2012; Gronlund & Neuschatz, in press; for an overview see Wixted, Gronlund, & Mickes, 2014). Mickes et al. (see also Clark, Erickson, & Breneman, 2011) showed that ratio-based probative value measures, which were assumed to be a measure of discriminability, are affected by response bias (a witness's willingness to make a response). Consequently, as response bias becomes more conservative, these ratio-based measures increase despite the ability of the witness to make a discrimination remaining unchanged. Therefore, ratio-based measures are inappropriate for comparing the performance of different identification procedures, especially if those procedures differ in the response bias they engender. It has been argued that showups result in a more liberal choosing rate (Meissner, Tredoux, Parker, & MacLin, 2005); therefore, it is possible that the liberal choosing may be driving lineup superiority.

Recently, researchers have demonstrated the utility of ROC curves when comparing eyewitness identification procedures (Mickes et al., 2012; Wixted et al., 2014; Gronlund et al., 2014). Despite researchers often collecting confidence ratings at the time of an identification, these data are typically analyzed separately and presented collapsed over all levels of witness confidence. ROC curves present the correct and false identification rates at each level of witness confidence, accumulated from the highest to the lowest confidence levels. Note that filler identifications are ignored, which means that the ROC curve does not sweep over the entire range of the ROC space (0 to 1) because only suspect identifications are utilized when creating the curves. The identification procedure that produces the best discriminability between the guilty and the innocent suspect yields the highest ROC curve (closest to the upper left-hand corner of the space; see Figure 1). To evaluate these differences statistically, the area under the

curve or partial area under the curve (AUC or pAUC) is computed for each identification procedure. The present study utilizes ROC analyses to evaluate identification performance at varying retention intervals and differing lineup fairness.

Hypotheses

Given the current state of showup research several hypotheses are tested. (1) Identification procedures conducted immediately will result in better performance (greater pAUC) than procedures conducted after a delay. (2) Immediate showups, when compared to delayed lineups, will result in equivalent or perhaps better performance. (3) Based on Wells and Quinlivan (2009), showups will result in superior performance compared to biased lineups, but fair lineups will result in better performance than showups. The suspect will always stand out in a showup or biased lineup because either the suspect is the only one present or the suspect stands out because the fillers are not reasonable options. In a fair lineup, the suspect would not stand out relative to the other fillers and should provide greater protection for the innocent suspect.

Method

Participants

A total of 1584 participants were recruited from the University of Alabama in Huntsville (*n* = 643), Canisius College (*n* = 194), University of Oklahoma (*n* = 212), Texas A&M University – Commerce (*n* = 152), Florida Southern College (*n* = 70), and SurveyMonkey (*n* = 293). Twenty additional participants were recruited by other means[2]. There were 506 males and 1078 females with a mean age of 24.8 years. For those participants who chose to indicate their race, the majority were Caucasian (69.6%), followed by African American (14.1%), Asian

[2] The UAHuntsville research laboratory recruited individuals to participate for course credit. Participants were randomly assigned to conditions in the same manner as all other participants and the results were not significantly different when excluded from analysis.

(6.4%), American Indian (1.7%), or other (8.3%). The vast majority of the participants received course credit in exchange for their participation. All participants were treated in accord with the ethical guidelines of the American Psychological Association and the Institutional Review Boards of the appropriate higher education institution.

Design

The current experiment conforms to a 2 (retention interval: immediate or delayed) x 3 (suspect: guilty, or innocentstrong or innocentweak, as defined below) x 3 (identification task: showup, fair lineup, or biased lineup) between-participants factorial design. The primary dependent variables are correct identifications of the guilty suspect, false identifications of the innocent suspects (innocentstrong or innocentweak), which were used in the calculation of pAUC's and confidence ratings.

Materials

Video. We utilized the Gronlund et al. (2009) mock crime video, which lasts approximately 1 min 45 s. It begins with a couple walking down a sidewalk. The male actor is seen entering a theatre as the woman continues walking across a street toward a parking lot. Interspersed amongst this footage are shots of a suspicious-looking man getting out of his car and hiding behind some bushes. As the woman approaches the parking lot, the suspicious man suddenly jumps out from the bushes, steals her purse, and runs off. The perpetrator's face is in view for approximately 15 seconds during the entirety of the video.

The Suspects. The study utilized two different innocent suspects who yielded different false identification rates in a prior study (Gronlund et al., 2009). Gronlund et al. referred to the frequently-chosen innocent suspect as "innocentstrong" and the rarely-chosen innocent suspect as "innocentweak", and we continue with these labels. The guilty suspect photo was taken

shortly after the mock crime. The suspect wore different clothing. It was the photo Gronlund et al. referred to as "guiltystrong".

Showups and Lineups. There was a guilty suspect and two innocent suspects for each identification task resulting in nine lineups and three showups. There were three lineups created for the guilty suspect and for each of the two innocent suspects: a fair and a biased lineup that also were utilized by Gronlund et al. (2009). The three showups included a photo of guiltystrong, a photo of innocentstrong, and a photo of innocentweak. For details regarding lineup composition of the fair and biased lineups, see Gronlund et al. (2009).

Procedure

The procedure closely replicated Gronlund et al. (2012) and was presented using online software (www.surveymonkey.com) in the laboratory. In the laboratory, participants were randomly assigned to conditions; participants who participated outside the laboratory were assigned to conditions based on date of birth (odd or even day and month). Online presentation in the laboratory ensured that the experimenters were blind to the identification conditions. After informed consent, participants viewed the video, which they were instructed to watch closely because they would have to answer questions about it afterwards. Participants in the immediate condition proceeded to the distractor task that consisted of solving 20 anagrams of U.S. state names (e.g., AALABMA). After the distractor task, participants were read unbiased instructions (the perpetrator from the video may or may not be present in the identification procedure) while they read along on the computer screen. After the instructions, participants proceeded to the identification task. In the showup conditions, participants viewed a single photograph and were instructed to indicate if the perpetrator was present or absent. Participants in the lineup conditions viewed six photographs simultaneously (two rows of three individuals) and were

asked to identify one of the six photographs as the perpetrator or indicate that the perpetrator was 'Not There'. After completing the identification task, participants rated their confidence on a 7-point scale anchored at 1 with "not at all confident" and at 7 with "extremely confident." Following the confidence rating, participants answered three questions about the video: Did a crime occur in the video? What was the item stolen in the video? What was the last item seen in the video? These questions served as a manipulation check. No participants were excluded based on these questions. In total, the immediate procedure took approximately 5 minutes to complete.

Participants in the delayed condition were excused from the first session after viewing the video with a warning to not discuss what was seen in the video. The participants returned 48 hours later to complete the distractor task and identification phase of the study exactly as the immediate condition participants had done. Overall only 37 participants or 6% did not return for second session.

Results

Showups and lineups were compared over different retention intervals (approximately 5 minutes versus 48 hours), lineup fairness (fair and biased), and innocent suspect resemblance to the perpetrator (innocentstrong and innocentweak). Patterns from the data did not differ between laboratory (n = 1271) and non-laboratory (n = 313) samples, and we collapse across these when describing the results. Traditional methods of comparing identification procedures are examined first followed by ROC analysis.

Correct and False Identification Rates

The suspect, foil, and rejection rates for showups and lineups can be found in Tables 1 and 2, respectively. Overall, showups yielded a correct identification rate of 59% and a false

identification rate of 49% for innocentstrong and 40% for innocentweak. The correct identification rate for showups conducted immediately was 64.8% and the false identification rate was 48%; however, after the 48-hour delay, both the correct identification and false identification rates dropped slightly to 55% and 41%, respectively. The correct identification rate for the lineups collapsed across fairness was 76% with a false identification rate of 37%. In the immediate lineup condition with a correct identification rate of 75% and the false identification rate of 37%, with little change to the delayed lineup conditions (71% and 37%, respectively).

Logistic Regression

Correct Identifications. Logistic regression was applied to correct identifications using retention interval (immediate versus delay) and identification task (showup, biased lineup, fair lineup) as predictors. All *p*-values were two-tailed unless otherwise noted; we adopted a .01 level of significance to protect against inflation of Type I error rates. These predictors reliably discriminated between correct identifications and all other decisions (rejections, foil identifications), X^2 (3, N = 423) = 12.38, p = .006. Retention interval did not affect correct identifications, and it did not interact with identification task. There was, however, an effect of presentation type, *Wald* X^2 (2, N = 423) = 11.57, p = .003, and individual chi-squares were conducted to break down this effect. We collapsed across retention interval because it had no effect on correct identifications. The showup never produced more correct identifications compared to either lineup type. It was equivalent to the fair lineup (X^2 (1, N = 281) = 2.60, p = .14, Φ = .096), and the biased lineup produced more correct identifications (X^2 (1, N = 277) = 11.56, p = .001, Φ = .204). There was no difference in correct identifications between fair and biased lineups, X^2 (1, N = 288) = 3.44, p = .08, Φ = .11).

False Identifications. We separately considered the false identification rates for the two innocent suspects: innocentstrong and innocentweak. As expected, innocentstrong was chosen more often (53% versus 33%), X^2 (1, N = 867) = 47.71, p < .001, Φ = .24. Retention interval and identification task did not discriminate between false identifications of innocentstrong and all other decisions, X^2 (3, N = 229) = 7.79, p = .05. By itself, retention interval had no effect on innocentstrong false identifications, and it did not interact with identification task, *Wald* X^2 (2, N = 229) = 7.33, p = .03. Consequently, we will not consider these data further. We also will not construct ROC curves because we have insufficient data involving innocentstrong.

A different pattern of results emerged for the innocentweak suspect. Retention interval and identification task together discriminated between false identifications and all other decisions, X^2 (3, N = 638) = 52.45, p < .001. Retention interval, again, had no effect on its own, nor did it interact with identification task. But there was an effect of identification task, *Wald* X^2 (2, N = 638) = 44.85, p < .001. After collapsing over retention interval, individual chi-squares by identification task revealed that showups yielded more false identifications than both fair (X^2 (1, N = 443) = 52.93, p < .001, Φ = .35) and biased lineups (X^2 (1, N = 430) = 6.73, p = .01, Φ = .13). Not surprisingly, there were more false identifications from the biased compared to the fair lineup, X^2 (1, N = 403) = 22.22, p < .001, Φ = .24.

Showups do not compare favorably with lineups. The showup never resulted in eyewitnesses choosing the guilty suspect more often. As for the protection of innocent suspects, the showup also was not preferred, as it yielded more false identifications than either lineup type. However, as mentioned earlier, separately assessing the correct and false identifications rates can be problematic (Wixted & Mickes, 2012), which is why ROC analyses are conducted to evaluate these identification procedures.

ROC Analysis

ROC curves comparing showups and lineups overall, as a function of retention interval, and as a function of lineup fairness, are all presented in Figure 1. Each ROC curve is derived from the correct and false identification rates at each confidence level and summarized by a trendline to better depict each curve. Each graph also includes the diagonal line where the correct identification rate equals the false identification rate, which represents chance performance. Note, as previously mentioned, that the *x*-axis extends from 0 to .60 because lineup ROC's are not traced out over the full range from 0 to 1 because only suspect identifications are utilized when creating the ROC curves. Consequently, pAUC values were computed using a false identification rate range from 0 to *q*, where *q* is set to a value slightly greater than the maximum false identification rate for the ROC's used in a comparison (see Wixted & Mickes, 2012).

The top panel of Figure 1 displays the ROC curves for the two identification procedures collapsed over retention interval and lineup fairness. There are two important points to highlight. First, it is clear that both lineups and showups produce discriminability above chance. Second, there is a significant difference in pAUC between showups (.16) and lineups (.29), $D = -5.84$, $p < .001$[3]. This pattern of results demonstrates that lineups result in better discriminability than showups, in replication of Gronlund et al. (2012).

Most interesting forensically, participants' performance in the delayed lineup conditions exceeded the performance of participants in the immediate showup condition (middle panel of Figure 1). The immediate showup (pAUC = .18) was significantly worse than the delayed lineup (pAUC = .27), $D = -3.31$, $p < .001$. Contrary to what some researchers and the criminal justice

[3] *D* is defined as (AUC1 – AUC2)/*s*, where *s* is the standard error of the difference between the two pAUCs. The standard error is estimated by the bootstrap method using 10,000 bootstraps (see Mickes et al., 2012; for a tutorial, see Gronlund et al., 2014).

system may have anticipated, these data indicate that an immediate showup did not result in better performance than conducting a lineup later, even when the lineup is delayed 48 hours. In fact, performance in the lineup conditions always exceeded performance in the showup conditions. The immediate lineup (pAUC = .30) was significantly better than the immediate showup (pAUC = .18), D = -4.30, p < .001, the delayed lineup (pAUC = .27) was significantly better than the delayed showup (pAUC = .14), D = -3.68, p < .001, and the delayed showup was significantly worse than the immediate lineup pAUC, D = 4.31, p < .001. In sum, a lineup identification, either immediately or at a delay, was more diagnostic than a showup. Neither of the lineup conditions were significantly different from one another as a function of retention interval, D = 1.05, p = .29, nor were the showup conditions, D = 1.03, p = .30..

The bottom panel of Figure 1 displays the ROC curves for the fair lineup, biased lineup, and showup. The fair lineup (pAUC = .31) and biased lineup (pAUC = .28) were significantly better than the showup (pAUC = .17), D = -6.13, p < .001 and D = -4.51, p < .001, respectively. Contrary to our hypothesis, there was no significant difference between the fair and biased lineup pAUC's, D = 1.22, p = .22 (for a similar result see Gonzalez et al., 1993, Experiment 2). However, false identifications for the fair lineup only extend to .11, whereas the biased lineup false identifications continue out to .26. That means that participants were more willing to make identifications at lower levels of confidence when the lineup was biased. Moreover, if limited to a forensically-relevant range (i.e., highest confidence decisions), the fair lineup ROC exceeds the biased lineup ROC. If we set q equal to .10, pAUC for fair was .04 and pAUC for biased is .03. The difference was not significant (D =1.25, p = .21), but if replicated, would signal that fair lineups are preferred for a confidence range that includes confident, accurate eyewitnesses from actual criminal cases.

Discussion

The present research resulted in a number of interesting and important findings. Most importantly, lineups were more diagnostic than showups. This was true when showups were compared to lineups immediately and after a 48-hour delay. These results replicate previous work comparing these identification procedures (Gronlund et al., 2012), and extend the findings to a longer retention interval. Although there was not a significant effect of retention interval over our 48-hour retention interval, the means did fall in the predicted direction. Additionally, an effect of retention interval may have been found given a longer interval (e.g., a week). It also was true that both fair and biased lineups resulted in better diagnostic accuracy than showups, contrary to the predictions of Wells and Quinlivan (2009).

There are several possible explanations for why lineups are a more diagnostic procedure than showups. One explanation involves the number of options at test, which could induce participants to adopt a different decision criterion across procedures (Gonzalez et al., 1993; Meissner, Tredoux, Parker, & MacLin, 2005). If presenting only one option created a more liberal criterion, increased choosing would lead to more correct and false identifications in the showup conditions. However, the results from the present experiment are inconsistent with this explanation, as showups yielded the lowest correct identification rate and the highest false identification (less diagnostic decisions). The ROC analyses, in this study and in Gronlund et al. (2012), demonstrate that a more liberal criterion in showups fails to explain the data.

An alternative explanation is that showups and lineups engage different decision processes. Gonzalez et al. (1993) argued that participants might invoke a two-stage process in a lineup. Participants first discover which person is the "best match," and then decide if that best match is the perpetrator. In showups, however, participants need only engage this second stage

because there is only one option from which to choose. This idea is similar to the relative and absolute distinction proposed by Wells (1984) and others (Cutler & Penrod, 1988; Lindsay & Wells, 1985). One problem with this explanation is that the processes involved in relative and absolute processing are not well specified (see Clark & Gronlund, in press). Gonzalez et al. (1993) also argued that a showup could be interpreted as a lineup with a functional size of one (i.e., a biased lineup). When there is only one viable option in a lineup, participants do not need to decide which option is the best match to the perpetrator. Consequently, the processing strategy would be similar for showups and biased lineups. However, the Gonzalez et al. results were not consistent with this explanation, nor were the present data.

Wixted and Mickes (2014) recently proposed a diagnostic-feature signal-detection-based theory that provides an explanation of why lineup performance should be superior to showup performance. When lineups foils are selected based on the characteristics of the perpetrator or suspect (see Clark, 2012), then all the foils should contain those characteristics. However, some features that are specific to the perpetrator will remain unique, as they were not used in the creation of the lineup. These unique features then become the diagnostic features when the faces are compared in a lineup. The theory proposes that better discriminability occurs in lineups because multiple lineup members can be compared. This allows diagnostic and non-diagnostic features to be distinguished, and the diagnostic features to subsequently receive more attention. A showup does not allow this comparison, and consequently, diagnostic features may never become apparent.

Limitations

Other factors undoubtedly affect the performance differences between these two identification procedures. For example, the current study employed unbiased instructions. Dysart

and Lindsay (2007) suggested that the demand to choose from a showup is greater in the field shortly after the crime, more so than when a lineup is conducted later. This could hamper showup performance. Conversely, showups in the field are likely to be performed live (see Valentine et al., 2012), and live showups may provide more retrieval cues at the time of identification. Research on lineups has shown that more realistic modes of presentation (videotape versus live) result in improved performance (Shapiro & Penrod, 1986; Cutler, Behrman, Penrod & Fisher, 1994). Melara, DeWitt-Richardson, and O'Brien (1986) found that correct identifications were higher when voices were added to the photographic lineup procedures. Consequently, because a live showup provides more retrieval cues to the witness, it could improve performance relative to a photographic lineup.

The lack of a significant effect of retention interval within identification type was unexpected, however, research on the effect of delay on identification performance has been inconsistent (see Dysart & Lindsay, 2007). Valentine et al. (2012) found a trend for participants to choose more from lineups when they were conducted after seven versus 28 days but retention interval had no affect on correct and false identification rates. In addition, Dysart and Lindsay (2007) concluded that the small body of literature on showups conducted after a delay failed to indicate a clear detrimental effect of delay on showup performance. The current data do not demonstrate a significant detrimental effect of delay within identification procedure at a 48-hour delay but at a further delay (e.g., one week) it is likely that this difference would be significantly greater given the pAUC's were greater for the immediate conditions than the delayed. Furthermore, ROC's allowed for the examination of discriminability differences between identification procedures at a delay. Future research should continue to examine the effects of retention interval on identification techniques using ROC analysis.

Practical Application

These data indicate that lineups are a more diagnostic procedure than showups. However, it is likely that at longer retention intervals (> 48 hours), lineup performance would eventually decline below that of showups conducted shortly after a crime. More research will be needed to determine just how long police have to create a lineup before the benefits of doing so are exhausted. Nevertheless, our point is that if the police have a suspect and want to conduct a showup, it may be preferable – from a memory perspective – to conduct a lineup even if it requires some additional time (up to 48 hours in the present study) to create it. The memorial benefits of immediate testing with a showup do not outweigh the potential costs of potentially prosecuting an innocent suspect. It is important to note that with technology today, police can create photographic lineups quickly. Given the unreliable nature of showups, we suggest that creating a lineup that follows best practice guidelines offers the most diagnostic procedure, and the best chance at protecting the innocent and implicating the guilty.

References

Beal, C. R., Schmitt, K. L., & Dekle, D. J. (1995). Eyewitness identification of children: Effects of absolute judgments, nonverbal response options, and event encoding. Law and Human Behavior, 19(2), 197-216. doi:10.1007/BF01499325

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*(5), 475-491. doi:10.1023/A:1012840831846

Bradley v. State, 264 SD.E.2d 332 (Ga. App. 1980).

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(2), 118-128. doi:10.1037/1076-898X.14.2.118

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology*, *17*(6), 629-654. doi: 10.1002/acp.891.

Clark, S. E. (2008). The importance (necessity) of computational modelling for eyewitness identification research. *Applied Cognitive Psychology*, *22*(6), 803-813. doi: 10.1002/acp.1484.

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

Clark, S. E., Erickson, M. A., & Breneman, J. (2011). Probative value of absolute and relative judgments in eyewitness identification. *Law and Human Behavior*, *35*(5), 364–380. doi:10.1007/s10979-010-9245-1

Clark, S. E., & Godfrey, R. D. (2009). Eyewitness identification evidence and innocence risk. *Psychonomic Bulletin & Review*, *16*(1), 22-42. doi:10.3758/PBR.16.1.22

Commonwealth v. Carter, 414 A.2d 369 (Pa. Super. Ct. 1979).

Cutler, B. L., Berman, G. L., Penrod, S., & Fisher, R. P. (1994). Conceptual, practical, and empirical issues associated with eyewitness identification test media. In D. F. Ross, J. D. Read & M. P. Toglia (Eds.), Adult eyewitness testimony: current trends and development (pp. 163-181). New York. Cambridge University Press.

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

Dekle, D. J. (2006). Viewing composite sketches: lineups and showups compared. *Applied Cognitive Psychology, 20,* 383-395. doi:10.1002/acp.1185

Dekle, D. J. (1997). Testing delays resulting in increase identification accuracy in lineups and showups. *Journal of Offender Rehabilitation, 25*, 35 – 49.

Dekle, D., Beal, C. R., Elliot, R., & Huneycutt, D. (1996). Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology, 10,* 1-12.

Dunning, D., & Stern, L. B. (1994). Distinguishing accurate from inaccurate eyewitness identifications via inquiries about decision processes. *Journal of Personality and Social Psychology, 67,* 818-835. doi:10.1037/0022-3514.67.5.818

Dunning, D., & Perretta, S. F. (2002). Automaticity and eyewitness accuracy: A 10- to 12- second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology, 87,* 951-962. doi:10.1037/0021-9010.87.5.951

Dysart, J. E., & Lindsay, R. C. L. (2007). Show-up identifications: Suggestive technique or reliable method? In Rod C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.),

The handbook of eyewitness psychology, Vol II: Memory for people. (pp. 137-153). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Flowe, H. D., & Ebbesen, E. B. (2007). The effect of lineup member similarity on recognition accuracy in simultaneous and sequential lineups. *Law and Human Behavior*, *31*, 33-52. doi:10.1007/s10970-006-9045-9

Flowe, H., Ebbesen, E., Burke, C., & Chivabunditt, P. (2001, July). At the scene of the crime: An examination of the external validity of published studies on line-up identification accuracy. Paper presented at the American Psychological Society Conference, Toronto, Canada.

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

Gonzalez, R., Ellsworth, P. C., & Pembroke, M. (1993). Response biases in lineups and showups. *Journal of Personality and Social Psychology*, *64*(4), 525-537. doi:10.1037/0022-3514.64.4.525

Goodsell, C. A., Gronlund, S. D., & Carlson, C. A. (2010). Exploring the sequential lineup advantage using WITNESS. *Law and human behavior*, *34*(6), 445-459. doi: 10.1007/s10979-009-9215-7

Goodsell, C., Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2013). Showups. B. L. Cutler (Ed.), *In Reform of Eyewitness Identification Procedures* (pp. 45-63). Washington, DC US: American Psychological Association. doi:10.1037/14094-007

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*(2), 140-152. doi:10.1037/a0015082

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

Gronlund, S. D., & Neuschatz, J. S. (in press). Eyewitness identification discriminability: ROC analysis versus logistic regression. *Journal of Applied Research in Memory and Cognition.*

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using ROC analysis. *Current Directions in Psychological Science, 23.*

Innocence Project. (2013). Understand the causes: eyewitness misidentification. Retrieved from http://www.innocenceproject.org/understand/eyewitness-misidentification.php on April 2nd, 2013.

Lane, S. M., & Meissner, C. A. (2008). A 'middle road' approach to bridging the basic-applied divide in eyewitness identification research. *Applied Cognitive Psychology, 22,* 779-787. doi:10.1002/acp.1482

Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21*(4), 391-404. doi:10.1023/A:1024807202926

Luus, E. C. A., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15*(1), 43 – 57. doi:10.1007/BF01044829

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

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33*(5), 783-792. doi: 10.3758/BF03193074

Melara, R. D., Dewitt-Rickards, T. S., & 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

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 4,* 361-376. doi:10.1037/a0030609

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

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 27*(5), 523-540. doi:10.1023/A:1025438223608

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

SurveyMonkey Inc. (2013). [Online computer software]. Palo Alto, California.

United States v. Wade, 388 U.S. 218 (1967).

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

Wagenaar, W. A., & Veefkind, N. (1992). Comparison of one-person and many-person lineups: A warning against unsafe practices. In F. Lösel, D. Bender, & T. Bliesener (Eds.),

Psychology and law: International perspectives. (pp. 275-285). Oxford England: Walter De Gruyter.

Weber, N., Brewer, N., Wells, G. L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: the unruly 10-12-second rule. *Journal of Experimental Psychology: Applied, 10,* 139-147. doi:10.1037/1076-898X.10.3.139

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

Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior,33,* 1-24. doi: 10.1007/s10979-008-9130-3

Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives on Psychological Science, 7*(3), 275–278. doi:10.1177/1745691612442906

Wixted, J. T. & Mickes, L. (in press). A signal-detection-based diagnostic-feature model of eyewitness identification. *Psychological Review*.

Yarmey, A. D., Yarmey, M. J., & Yarmey, A. L. (1996). Accuracy of eyewitness identification in showups and lineups. *Law and Human Behavior, 20*(4), 459-477. doi:10.1007/BF0149898

Table 1

Percent Chosen for Showup Identification Procedures

	Perpetrator	Innocentweak	Innocentstrong
Immediate			
Suspect	62.1	42.0	53.8
Filler	N/A	N/A	N/A
Rejection	37.9	58.0	46.2
N	74	193	39
Delay			
Suspect	55.7	38.1	43.5
Filler	N/A	N/A	N/A
Rejection	44.3	61.9	56.5
N	61	42	46

Table 2

Percent Chosen for Lineup Identification Procedures

	Perpetrator		Innocentweak		Innocentstrong	
	Fair	Biased	Fair	Biased	Fair	Biased
Immediate						
Suspect	68.3	81.4	10.2	28.1	46.3	64.9
Filler	10.0	2.90	54.2	31.6	24.4	2.70
Rejection	21.7	15.7	35.6	40.3	29.3	32.4
N	60	70	118	114	41	37
Delay						
Suspect	68.6	75.0	11.1	28.7	41.4	65.8
Filler	15.1	5.60	51.1	31.3	44.8	15.8
Rejection	16.3	19.4	37.7	40.0	13.8	18.4
N	86	72	90	81	29	38

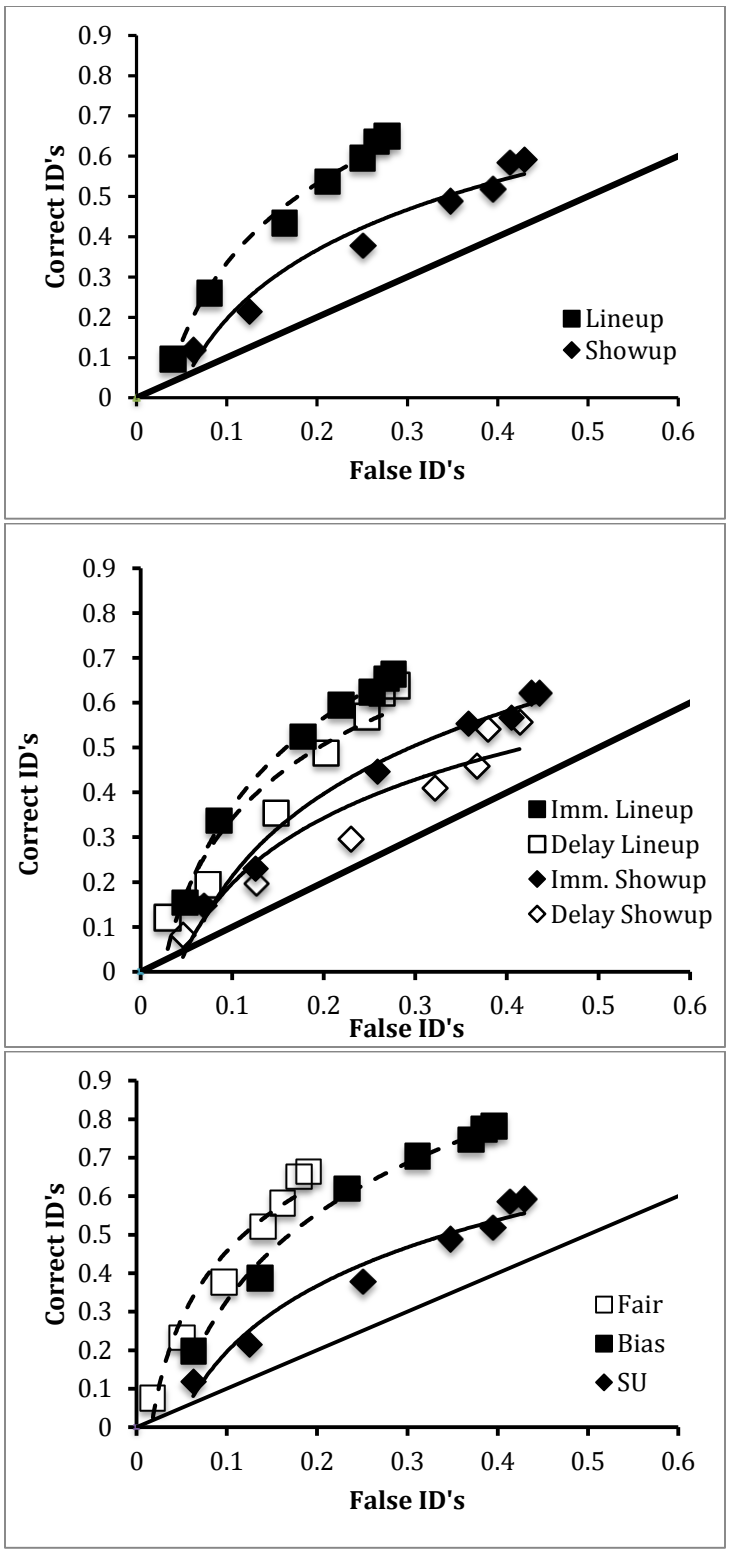

Figure 1. ROC curves comparing showups (SU) and lineups (LU; *top panel)*, showups and lineups at the different retention intervals (*middle panel)*, and the lineups separated by fairness *(bottom panel)*. The thick diagonal line in each panel indicates chance performance.