Loughborough University

Loughborough University
Institutional Repository

The effects of target discriminability and criterion placement on accuracy rates in sequential and simultaneous target-present lineups

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: FLOWE, H.D. and BESSEMER, A., 2011. The effects of target discriminability and criterion placement on accuracy rates in sequential and simultaneous target-present lineups. Psychology, Crime & Law, 17(7), pp. 587-610.

Additional Information:

- This is an Accepted Manuscript of an article published by Taylor & Francis in Psychology, Crime and Law on 2011, available online: http://dx.doi.org/10.1080/10683160903397540.

Metadata Record: https://dspace.lboro.ac.uk/2134/20194

Version: Accepted for publication

Publisher: © Taylor & Francis

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.

Running head: Target Discriminability and Criterion Placement Effects in Lineups

The Effects of Target Discriminability and Criterion Placement on Accuracy Rates in Sequential and Simultaneous Target Present Lineups

Heather D. Flowe
University of Leicester

Anneka Bessemer
University of California, San Diego

DRAFT

Address ms. correspondence to:

Heather D. Flowe
hf49@le.ac.uk
School of Psychology
Forensic Section
University of Leicester
106 New Walk
Leicester, UK
LE1 7EA

Abstract

Participants attempted to select previously studied faces from lineups that were administered either sequentially (test faces presented one at a time) or simultaneously (test faces presented altogether). Target discriminability was manipulated by varying the number of facial features that could be used to distinguish the study face from the other test faces (foils) or by varying study face exposure duration. In addition, decision criterion level was manipulated via an instruction manipulation. Results indicated that sequential participants adopted a stricter decision standard under high criterion instructions compared to simultaneous participants. Under liberal criterion setting instructions, the rate at which the target was selected was comparable across lineup procedures. Target discriminability affected target selections to a greater extent in simultaneous compared to sequential lineups. We discuss the applied implications that these findings have regarding correct identifications from lineups.

DRAFT

A Comparison of Decision Processes in Sequential and Simultaneous Target Present Lineups

Psychologists have developed a number of procedures that reduce the rate of mistaken identifications of criminal lineups, including lineup construction techniques that affect the discriminability of the suspect relative to the other lineup members (e.g., Malpass, 1981; Wells, Rydell, & Seelau, 1993), cautionary instructions that warn eyewitnesses that the culprit may not be present in the lineup (e.g., Clark, 2005; Malpass & Devine, 1981; Steblay, 1997), and lineup presentation procedures (e.g., Lindsay & Wells, 1985; Steblay et al. 2003). These procedures, however, might also reduce the rate of correct positive identifications (Clark, 2005). In this study, we examined the combined effects of target discriminability and cautionary instructions on correct positive identifications in two different types of lineup procedures; namely, sequential and simultaneous lineups.

In a traditional simultaneous photographic lineup procedure, all of the test faces are displayed simultaneously in an array. The witness then makes a single identification response and either identifies a face as that of the culprit, or rejects the lineup if none of the faces are sufficiently similar to the culprit in memory. In a sequential lineup procedure, the faces are presented one at a time, and the witness makes a yes/no decision for every face. The witness may respond "yes" to a particular face, and thereby positively identify that face as the perpetrator, or the witness may indicate "no" to every face, and thereby reject the lineup because none of the faces is sufficiently similar to the perpetrator in memory. Once an eyewitness indicates "yes" to a face, the administrator might conclude the lineup test, even if there are additional faces to present, or else allow the eyewitness to view the remaining photographs (see McQuiston-Surrett, Malpass, & Tredoux, 2006). Another procedural difference between simultaneous and sequential lineups

DRAFT

is that the sequential witness does not know in advance how many test faces will be displayed. This procedural detail is included to prevent the witness from identifying the last face in the series simply because no other faces will be shown.

Fewer false alarms (i.e., the rate at which innocent suspects are identified) occur in a sequential compared to a simultaneous lineup, particularly if the lineup is biased (Carlson, Gronlund, & Clark, 2008). The decrease in false alarm rates found for sequential compared to simultaneous lineups has been referred to in the literature as the *sequential superiority effect*, and sequential lineup procedures have been proposed as an alternative to simultaneous lineups (TWGEE, 1999). In simultaneous lineups, it has been theorized that false alarms occur more often because a relative decision strategy is operative, wherein participants identify a face if it is a relatively better match to the culprit represented in memory compared to the other faces in the array (Lindsay & Wells, 1985). In contrast, since participants tested with sequential lineups cannot directly compare the test faces because they are being presented one at a time, it has been theorized that witnesses have to rely more on an absolute decision strategy, wherein each test face is compared to the witness' memory representation of the culprit. Several studies (e.g., Cutler & Penrod, 1988; Lindsay, Lea, Nosworthy, & Fulford, 1991; Sporer, 1994; Wells, 1993) have been interpreted as being in line with predictions derived from Lindsay and Wells' (1985) relative versus absolute judgment theory (for a review, see Wells et al. 1998).

Since the Lindsay and Wells (1985) seminal study, additional research comparing simultaneous and sequential lineups has been conducted (for a recent discussion of this work, see Malpass, Tredoux, & McQuiston-Surrett, 2009, and Lindsay, Mansour, Beaudry, Leach, & Bertrand, 2009). Results suggest that sequential procedures should be used in practice instead of simultaneous lineups because mistaken eyewitness identification rates would be reduced without

DRAFT

appreciably affecting the hit rate, or the rate at which guilty suspects are identified. Meta-analyses, which is a statistical method that aggregates results across studies and thereby provides greater power to detect small effects that might not otherwise be observed in a single study, of the lineup literature has indicated that not only is the false alarm rate lower in sequential lineups, but there is also a small but reliable effect on the hit rate, such that correct positive identification rate is lower for sequential compared to simultaneous lineup presentations (Steblay et al., 2001; Ebbesen & Flowe, 2002; cf. Clark, Howell, & Davey, 2008). The small but reliable effect of lineup procedure on hit rates has indicated that additional theoretical detail regarding decision processes in lineups is needed. The intention of the present research is to advance our theoretical knowledge of target present simultaneous and sequential lineups so that steps can be taken to increase the rate of correct positive identifications in practice, a goal that other psychologists have also pursued (e.g., Levi, 1998).

In the present study, we investigated the conditions under which hit rates differ in simultaneous and sequential lineups. In particular, identification responses in target present lineups that were presented simultaneously or sequentially were compared under varying conditions of criterion placement and target discriminability. The first hypothesis tested was that a liberal criterion shift, which can be induced in practice by leading the witness to believe that the perpetrator is in the lineup, would eliminate differences between procedures in the hit rate. Second, we hypothesized that the difference in the hit rate across lineup procedures would vary in relation to target discriminability. Target discriminability refers to the degree to which the target matches the study face (or perpetrator) in memory. Target discriminability can be affected in a number of ways, including manipulating the similarity of the lineup target to the study face (Flowe & Ebbesen, 2007), or by varying the duration of exposure to the study face (Meissner et

DRAFT

al., 2005). As the target is made more similar to the study face or studied for a longer period, the probability that the test faces matches the memory representation of the study face should increase. As a consequence, it should become easier to discriminate the target from the foils as target similarity and/or study exposure time increases. In the present study, target discriminability was manipulated by varying the number of lineup target features that matched the study face and by varying study face exposure time. Hit rates were expected to increase to a greater extent in simultaneous compared to sequential lineups as target discriminability was increased. The basis for these predictions is described next.

The Hypothesized Effect of a Liberal Criterion Shift on Hit Rates in Sequential and Simultaneous Lineups

One account that has been put forward to account for the difference in hit rates (as well as false alarm rates) in sequential and simultaneous lineups is derived from a signal detection framework (Ebbesen & Flowe, 2002). According to this view, the primary difference between simultaneous and sequential lineups is that sequential witnesses apply a stricter decision standard at test compared to simultaneous witnesses. In a sequential presentation, the lineup unfolds one photograph at a time. Consequently, witnesses might adopt a stricter standard because for any one face that they are viewing, there may be a photograph that they have not seen yet that could be an even better absolute match to the culprit in memory. They may find that a face matches the culprit in memory, but decline to identify the face, considering that there may be a stronger match still to come. Such a decision process would lead to fewer lineup choices overall, and decrease not only the false alarm rate, but the hit rate as well. We will now review this account in a bit more detail so that we may describe how criterion shifts within lineup procedures might be expected to impact hit rates.

DRAFT

Figure 1 shows a signal detection representation of the hypothesized effects of lineup presentation procedure on hit rates in a target present lineup. One distribution represents the culprit and the other distribution represents the foils. The distributions are aligned on an evidence dimension, which represents evidence that a lineup test face matches a target in memory. The evidence dimension is assumed to derive from both familiarity and recollection memory processes (see Wixted, 2007). Figure 1 depicts that the target faces (i.e., perpetrators in real life cases, or faces that have been studied in a memory experiment) are higher in memory strength on average compared to foil faces. For simplicity, the variability of the target and foil distributions are presented as being equivalent, though in practice, the variability of the foil distribution may be affected by the similarity of the foils with respect to each other, the lineup suspect, as well as their similarity to the perpetrator (also see Meissner et al. 2005). Additionally, the variability of the target distribution can be affected by memory strength, with variability increasing as memory for the target is strengthened (Mickes, Wixted, & Wais, 2007; Ratcliff, Sheu, & Gronlund, 1992, also see Ebbesen & Flowe, 2002). Figure 1 also depicts criterion standards for simultaneous (M) and sequential (Q) lineups, with the latter criterion hypothesized as being set at a more conservative level. The criterion represents the minimum level of evidence required for a positive identification of a face. Faces that exceed the criterion will be positively identified. With a more conservative criterion, a smaller proportion of target and foil faces will be positively identified than if a more liberal criterion were adopted. In Figure 1, the area on the target distribution that falls between the simultaneous and sequential decision criteria represents the difference in the hit rate that is expected in this hypothetical example. As can be seen, the differential criterion placement account predicts that the hit rate will be larger in simultaneous compared to sequential lineups.

DRAFT

Meissner et al. (2005) empirically tested the criterion shift explanation of responses in simultaneous and sequential lineups that Ebbesen and Flowe (2002) put forth, and found that criterion placement in sequential lineups was set at a more conservative level compared to simultaneous lineups. Meissner et al. (2005) hypothesized that the operation of fluency-based recognition processes (see Yonelinas, 2002 for a review) in simultaneous lineups was affecting the decision standard applied at test. Perceptual fluency, or ease of processing, can be used as a memory cue in simultaneous lineups because the test faces can be directly compared. In line with their predictions, they found evidence using the remember-know procedure (Gardiner & Richardson-Klavehn, 2000) that simultaneous identifications may be influenced by fluency processes to a greater extent than sequential identifications. Additionally, the use of recollection processes did not seem to differ depending on whether the test faces were viewed simultaneously or sequentially. In contrast, Gronlund (2005) proposed that sequential identifications are based on recollection processes to a greater extent than are simultaneous identifications, and he found evidence consistent with his hypothesis in his research (see also Carlson, Gronlund, & Clark, 2008).

In the present project, we examined the effect of a liberal criterion shift on accuracy rates in sequential and simultaneous lineups. The effects of criterion placement were examined by manipulating participants' expectations of whether the target was present in the lineup. We surmised that under conditions in which participants are led to believe that the target may not be present in the lineup, differences in criterion placement across lineup procedures would evince, and the hit rate would be lower in sequential compared to simultaneous lineups. Psychologically, participants in a sequential lineup may withhold identifying a face unless it is a strong match to the study face in memory. Strong matches are preferred because the study face may not be

DRAFT

present in the lineup and because there may be a yet to be seen face in the queue that may be an even better match to the study face in memory. In contrast, we expected that if participants were led to believe that the target was present in the lineup, participants would set a more liberal decision standard in sequential lineups than otherwise would be the case. Compared to the high criterion condition, participants would be more inclined to identify faces with a weaker match because they have been told that the study face is present in the lineup. Hence, they will identify a weaker match because they will not want to miss the opportunity to identify the study face. As such, the rate of correct positive identifications will be comparable across lineup procedures under liberal criterion conditions.

The Hypothesized Effect of Target Discriminability on Hit Rates in Simultaneous and Sequential Lineups

The present study also extends previous research by comparing the target selection rate in simultaneous and sequential lineups at different levels of target discriminability. Signal detection analysis of the problem suggests that the effect of target discriminability on target selections (i.e., the hit rate) will differ depending on decision criterion placement. Specifically, increasing target discriminability will increase the hit rate in both lineup procedures, but the size of the effect will be larger in the simultaneous condition because the decision criterion is more liberal.

This prediction is demonstrated in Figure 2, which displays the receiving operator characteristics (ROCs) that are expected under four different target discriminability (d') conditions. ROC analysis allows for examining the independent contributions of discriminability and criterion placement on memory performance, which is measured by the hit rate, or the rate at which the target is identified (y-axis), and the false alarm rate, or the rate at which a foil (or an innocent suspect) is identified (x-axis). With higher discriminability levels, there is less overlap

DRAFT

between the target and foil distributions. Holding the decision standard constant, the target will be identified at a higher rate as discriminability increases. This effect is reflected in Figure 2 by the height of the curves with respect to the x = y diagonal line (which represents conditions in which the target and foil distributions completely overlap). Additionally, the points along each ROC curve shown in Figure 2 represent decision criterion placement (or decision confidence). Moving along the curve from left to right represents the adoption of an increasingly lenient criterion (i.e., less evidence that a test face is one of the targets in memory is needed for making a positive identification). As the criterion becomes more lenient (or decision confidence decreases), the positive identification rate increases, resulting in more hits and false alarms. The line that is perpendicular to the x = y diagonal represents the criterion cut point for an unbiased witness. Criteria that lie to the left of the line represent witnesses who are more conservative, and criteria that lie to the right represent witnesses who are more liberal. Assuming that the decision standard is relatively more lenient in a simultaneous compared to a sequential lineup, we can compare the effects of lineup procedure and discriminability on accuracy by comparing the hit and false alarm rates at different criterion cut points. In Figure 2, the criterion for simultaneous lineups is displayed with gray circles, and the sequential lineup criterion is displayed with black circles. The criteria levels for simultaneous and sequential lineups are held constant across discriminability levels. As shown in Figure 2, the effect on hits is expected to be larger in a simultaneous compared to a sequential lineup as d' is increased. Note that the size of the difference becomes smaller as one becomes extremely biased (i.e., under conditions in which the criteria fall to the right of the line demarcating an unbiased witness). In the present project, we tested these predictions by manipulating target-foil discriminability. Experiments 1 and 2 employed a feature identification task, whereby the discriminability of the target was varied by

DRAFT

manipulating the number of target features that matched the study face. Experiment 2 also employed an exposure duration manipulation; hence, target discriminability should theoretically increase as study face exposure duration increased.

Experiment 1

Method

Participants. A total of 206 undergraduates participated in exchange for course credit.

Materials. In order to manipulate target discriminability, we elected to utilize composite drawings as face stimuli because they enable an operational definition of target discriminability that follows a clear standard. Specifically, the use of composite faces enabled us to parametrically measure target discriminability by manipulating the number of facial features that the target and foils shared in common as well as the number of features that the lineup target shared with the study face. As the number of features added to the target face that matched the study face increased, both configural and featural similarity of the target face with respect to the study face increased, whereas the similarity of the target to the foil faces decreased. Our choice of stimuli was also motivated by the desire to begin with the simplest instance of a face, at least initially, to illuminate differences in simultaneous and sequential lineup decision processes. In the General Discussion we further discuss these issues.

An illustration of a study face and an accompanying set of lineups is presented in Figure 3. A stimulus pool of 18 study faces and accompanying lineup sets was created. Face stimuli were constructed using FACES 3.0, a composite drawing software program used by law enforcement. To compose a lineup set, first a single face was created using a function of the software that allows the user to generate a face at random. Facial hair and head hair were removed from the resulting face so as to focus participants' attention on the facial features, as previous research has

DRAFT

found that hair style and facial hair can reduce identification accuracy (Patterson & Baddeley, 1977; Read, Vokey, & Hammersley, 1990). This randomly generated face served as the *base face*, which was used to create the other faces for the lineup. For the lowest target discriminability level (level 2), the 6 base faces were manipulated such that each face had two features that differed, namely the eyebrows and noses, while all other features across the faces were identical. For the next highest discriminability level (level 3), one additional feature variation was added (eyes, mouth, or face), while all of the other remaining facial features were identical across faces. As for the two highest discriminability levels (levels 4 and 5), the remaining facial feature variations were added, such that 4 facial features varied across the faces and one feature was held constant across faces in feature level 4, and all 5 features across the test faces varied in feature level 5. For the feature level 3 and 4 lineups, lineups for all possible combinations of the order in which the specific facial feature types (eyes, mouth, face) could be introduced were constructed. Finally, one of the faces was selected from the feature level 5 lineup, and this face served as the study face for the set. In total, a stimulus pool of 864 faces, 180 lineups, and 18 study faces were constructed. As a result of the construction procedures, the lineup target became increasingly like the study face across feature levels 2-4, and the target was identical to the study face in the feature level 5 lineups. In addition, the foil faces and the target face became increasingly dissimilar across feature levels 2-5.

The position of the study face within the lineup was counterbalanced such that half of the study faces appeared early (in positions 2 and 3) in the lineup, and the other half appeared late (in positions 4 and 5). The study faces were 18 x 16 cm in size, whereas the test faces were 5.5 x 6.5 cm in size. The study and test faces were printed in grayscale on white paper. The study faces that were presented to a given participant were randomly determined. In addition, the particular

DRAFT

feature level presented in the lineup for a given study face was randomly determined across study faces, with the restriction that each participant see four lineups in total, one at each feature level.

Design and Procedure. Feature level, or the number of features that participants could use to identify the target (2, 3, 4, or all 5 features), was controlled within participants, and lineup procedure (simultaneous or sequential) and criterion level (low or high) were controlled between participants.

Participants studied 4 faces and were tested on 4 lineups, one for each study face. We ensured via visual inspection of the faces (and by comparing the feature codes that were generated for each by FACES) that that none of the faces shared any features in common. The study faces and lineup tests were administered to the participant by an experimenter who was not aware of the study hypotheses. Each participant studied four faces in succession for 10 s each, and then after a 5 min retention interval was tested on four individual lineups, one for each study face. Study face presentation duration and the retention interval duration were measured by the use of a stopwatch.

All participants were given practice lineups (one at every feature level) so as to demonstrate how the lineup faces would vary across feature levels. Participants were told to ignore any features that were identical across the faces. They were instructed to focus their attention on the features that differed across the faces, identifying the one face that matched the study face on the features that were varying. Participants were also correctly informed that each lineup was created with a particular study face in mind, and that the lineups would not necessarily be presented in the same order in which they had studied the faces. (The order in which the lineup feature levels were presented to the participant was randomized.)

DRAFT

Participants in the sequential condition viewed the lineup faces one at a time, and they were told to indicate for each face whether or not it was one of the faces that they had studied, whereas participants in the simultaneous condition were presented with all of the test faces and they were told to either to reject the lineup or to identify the study face by position number. Sequential participants could not see the remaining faces in the lineup (if any) once they had made a positive identification, nor could they review or change their mind about faces they had rejected previously. In the General Discussion, we will compare the procedures that were utilized to conduct sequential lineups in the current study with the sequential procedures that are recommended in practice (TWGEE, 1999), drawing attention to differences between the two that may be important for theoretical development as well as for applied practice.

With respect to the criterion manipulation, participants in the low criterion condition were instructed that the lineup "always contains a target face, and therefore you should always identify someone from the lineup". Participants in the high criterion condition were told that one of the study faces "may or may not be present in a given lineup, and therefore, you should refrain from choosing a face from a lineup if you do not see a face that matches one of the study faces on some features".

Once the participant made an identification decision, decision confidence was obtained using a 1-5 scale, with 1 representing "just a guess", and 5 representing "absolutely certain".

Measures. At every feature level, responses (a target or a foil face was selected, or no choice was made from the lineup) were tabulated for every participant. Choice data (rate at which any face was selected from the lineup), accuracy data (rate at which the target was selected from the lineup), and target selection/choice data (rate at which the target was selected for choosers) were

DRAFT

then aggregated across participants and analyzed as a function of feature level, lineup procedure and criterion condition.

With respect to measuring identification accuracy, it is important to point out that participants were essentially engaged in a feature identification test. When participants selected the target at feature levels 2-4, they were doing so on the basis of finding that the target had one or more features that matched the study face. At feature level 5, participants were engaged in an identification test, as selecting the target at this stage was based on determining that a test face was an identical match to a study face on all features. Therefore, the hit rate in the context of this experiment was defined as positive selections of the target that were made on the matching one or more target features to the study face in memory.

Target and foil selection rates were accumulated across confidence levels (starting from the highest confidence level and ending with the lowest) to construct ROC plots for an analysis of feature level by lineup procedure (see Yonelinas & Parks, 2007 for a review of ROC analysis in recognition memory research).

Results and Discussion

Preliminary Analyses. Preliminary analyses were performed to determine whether lineup target position and study face presentation order systematically affected the results. Sequential witnesses were aware that there were 6 lineup faces; hence, they may have been more likely to guess and select the last members of the lineup simply because they knew that no other faces were going to be presented. Results indicated that was not the case, as the rate of selecting faces across test faces 1-6 was nearly identical in simultaneous (.06, .20, .23, .10, .23, and .17) and sequential (.05, .20, .21, .11, .24, and .19) lineups. Next we examined whether the target

DRAFT

selection rate varied across the 4 test trials. Results indicated no significant effect of trial order on identification accuracy, $F < 1$ (*M*'s on test trials 1-4, respectively: .36, .35, .31, and .35).

Confidence ratings were significantly related to accuracy for choosers, both in simultaneous lineups ($r = .38$, $p < .01$) and in sequential lineups ($r = .32$, $p < .05$), a result that is in keeping with the Sporer et al. (1995) meta-analysis of the confidence-accuracy literature. Hence, the confidence data were further analyzed to assess the effect of a liberal criterion shift on accuracy rates in simultaneous and sequential lineups.

Statistical Approach. A linear mixed model analysis (Cnaan, Laird, & Slasor, 1997) was undertaken to test the study hypotheses because the response data across feature levels were unbalanced (i.e., subjects varied in whether they chose a face at a given feature level). The data were also analyzed with generalized estimating equations (see Hanley et al., 2003), which is an approach for analyzing repeated measures designs that controls for correlated errors across the response variable, and the results did not differ from the linear mixed model results. Therefore, only the results from the linear mixed model are reported for simplicity. Choice, accuracy, and target selection/choice data were analyzed separately, each with a full factorial analysis. Lineup procedure and criterion level were entered as the between subjects factors, and feature level was entered as the repeated measure.

Choosing. The choice rate analysis indicated significant main effects for lineup procedure, $F(1, 743.50) = 13.33$, $p < .01$, and criterion level, $F(1, 743.50) = 42.17$, $p < .01$. Participants were more likely to choose a face if the lineup was presented simultaneously rather than sequentially ($M = .89$ versus $M = .76$, respectively), or if they were in the low compared to high criterion condition ($M = .92$ versus $M = .71$, respectively). The effect obtained for lineup procedure is in line with the hypothesis that sequential participants adopt a more conservative judgment standard

DRAFT

than simultaneous participants. Additionally, the significant increase in choosing obtained for participants who were told that the study face was always present indicated that the criterion manipulation worked as intended. A significant effect on choosing was also obtained for feature level, $F(3, 460.17) = 9.08$, $p < .01$, but this effect was qualified by a significant interaction that was obtained for feature level and criterion level, $F(3, 451.24) = 5.43$, $p < .01$. The interaction arose because feature level did not affect choice rates when participants were told that the study face was always present (feature levels 2-5 *M*'s: .91, .93, .89, and .96), and it did affect choice rates when participants were told the study face may not be present (feature levels 2-5 *M*'s: .54, .66, .72, and .90). No other interaction effects were obtained.

Target Selections. The rate of target and foil choices by feature level, lineup procedure and criterion level are shown in Figure 4. With respect to target selections, results indicated a significant main effect for lineup procedure, $F(1, 792.42) = 3.89$, $p < .05$, and criterion level, $F(1, 792.42) = 20.05$, $p < .01$. As with the choice data, the target was selected at a higher rate in the simultaneous compared to the sequential condition ($M = .48$ versus $M = .41$, respectively), or if the decision criterion was set at a lower rather than higher level ($M = .53$ versus $M = .36$). Moreover, when participants expected that the target was present in the lineup, the target selection rate did not differ across lineup procedures (simultaneous $M = .54$, sequential $M = .51$). However, if participants were told that the target may not be present in the lineup, the target selection rate was significantly larger in the simultaneous compared to the sequential condition $F(1, 158) = 10.11$, $p < .01$, (simultaneous $M = .42$, sequential $M = .30$).

The next set of analyses addressed the hypothesis that target discriminability would have a larger effect on target selection rates in the simultaneous compared to the sequential condition. A significant effect on accuracy was obtained for feature level, $F(3, 390.05) = 37.63$, $p < .01$,

DRAFT

indicating that target selections increased as target discriminability increased (feature levels 2-5 *M*'s: .22., .35, .49, and .72). In the high criterion lineup instruction condition, feature level had a larger effect on accuracy in simultaneous (ß = .45) compared to sequential lineups (ß = .38), which was predicted by the signal detection analysis, but the interaction term for feature level and lineup procedure in the mixed model analysis was not statistically significant.

Accuracy may have been lower in the sequential condition either because sequential participants were less inclined to choose any face and hence, missed the target, or because they were more likely to identify foils. To examine this issue, we compared the rate at which the target versus a foil was selected across lineup procedures. The results indicated that criterion level, $F(1, 729.06) = 41.34$, $p < .01$, lineup procedure $F(1, 729.06) = 13.07$, $p < .01$, and feature level, $F(3, 451.24) = 8.90$, $p < .01$, significantly affected target selection/choice rates. In addition, criterion level had an interactive effect with feature level on target selection/choice rates, $F(3, 451.24) = 5.32$, $p < .01$. These results obtained because when the decision standard shifted to a lower level, the target was chosen more often than a foil face (simultaneous $M = .88$ versus sequential $M = .75$; low criterion $M = .92$ versus high criterion $M = .71$). With respect to the feature level results, the target was increasingly likely to be selected over a foil face as feature level increased in the high criterion condition (.54, .66, .72, and .90, across feature levels 2-5, respectively), whereas choosers selected the target more than 89% of the time at every feature level in the low criterion condition.

Confidence. In the final analysis, we turned to the confidence data to further examine the effects of liberal criterion shifts on target and foil selections. Accuracy rates within lineup condition varied with decision confidence in a manner that was consistent with the predictions derived from the signal detection analysis. As shown in the ROC's displayed in Figure 5, moving

DRAFT

from the high to more moderate confidence levels was associated with a bigger change in the target selection rate for the sequential compared to the simultaneous condition. The target and foil selection rates at each of the 5 confidence levels were z-transformed and entered into a regression equation to compare the slopes of the corresponding z-ROCs for the sequential and simultaneous conditions. This analysis allowed for a comparison of the rate of change in the target selection rate across lineup conditions as decision confidence decreased. The analysis indicated that the z-slope was significantly larger in the sequential compared to the simultaneous condition (.87 versus .79), $t(6) = 3.18$, $p < .05$. These results indicate that the target selection rate increased at a faster rate in the sequential compared to the simultaneous condition as the decision criterion moved to a more liberal level.

Summary of Findings. The rate of choosing any face was lower if the faces were presented one at a time rather than simultaneously, a result that is in line with previous theoretical analyses (Ebbesen & Flowe, 2002) and empirical findings (Meissner et al., 2005). Additionally, in the low criterion condition, in which the criterion was made comparable across lineup procedures, the target selection rate across lineup procedures was equal, whereas in the high criterion condition, in which the criterion was free to vary across lineup procedures, the target selection rate decreased in the sequential compared to simultaneous condition. These results support the hypothesis that a liberal criterion shift would have a larger effect on hit rates in sequential compared to simultaneous lineups. With respect to our hypothesis that target discriminability would affect hit rates to a larger extent in simultaneous compared to sequential lineups, the results did not support our prediction. Accuracy rates in both the simultaneous and sequential condition increased as target discriminability increased. Though the size of the target discriminability effect was larger in simultaneous compared to sequential lineups, the difference

DRAFT

was not statistically significant. Finally, the results further indicated that high criterion participants were being especially cautious when target discriminability was low, as they chose less often at lower compared to higher feature levels. In addition, participants who could have identified the target seemed to have been "weeded out" by the cautionary instructions that were delivered in the high criterion condition, especially at lower feature levels. Seemingly, participants with a high criterion rejected the lineups not because they did not recognize the target face, but rather because the target face did not exceed their criterion. This implies that target memory strength may have been lower for those who did make an identification in the high criterion condition. This interpretation is plausible because target selections significantly increased and foil identifications decreased when participants were encouraged to always select a face.

Overview of Experiment 2. The relationship between target discriminability and target selections was further explored in simultaneous and sequential lineups. Two methods were utilized to control target discriminability. In the first method, we borrowed from psychophysics the *method of ascending limits*, and presented participants with a study face followed by a series of lineups that differed with respect to target discriminability. The lineups were presented using a simultaneous or sequential method. As in Experiment 1, target discriminability was operationalized by varying the number of features that differed across the lineup faces. Lineups presented early in the series had few features differing across the lineup members compared to the lineups appearing later in the series. Participants in Experiment 2, however, were able to self-adjust target discriminability to the point at which they could select the target. We expected that if the lineup was presented sequentially rather than simultaneously, participants would request a larger number of features—in other words, more evidence that the target matches the study face

DRAFT

in memory, before making a positive identification. The logic is that more evidence consistent with memory should be required in order to surpass a relatively higher criterion. Therefore, sequential participants should require that a greater number of target features match the target in memory before making a positive identification compared to simultaneous participants.

The second method that was used to control target discriminability was the length of time that participants were able to examine the study face. Drawing from the signal detection framework presented earlier, we expected that duration of exposure would have a larger effect if the criterion were set at a lower rather than higher level. If true, accuracy in the simultaneous condition should be affected to a greater extent by study face exposure duration compared to the sequential condition. Duration of exposure (3 s versus 6 s) has been manipulated before in simultaneous and sequential lineups, and though it affected overall accuracy, no significant differences were found across lineup procedure (Meissner et al., 2005). Duration of exposure was manipulated in the present study by using durations that were appreciably higher (10 s versus 30 s) to further examine the effects of memory strength on accuracy in simultaneous and sequential lineups.

Experiment 2

Method

Participants. A total of 100 undergraduates participated to fulfill a course requirement.

Materials. The 18 study faces and attendant lineup sets from Experiment 1 served as the stimuli.

Design and Procedure. A 4 (feature level) x 2 (lineup procedure) x 2 (study face duration of exposure) mixed design was employed. Feature level was controlled within subjects. Participants studied a face and were presented with a series of 5 lineups (see Figure 3). The first lineup

DRAFT

presented faces that were identical on all features (i.e., they were clones). Participants were instructed to not pick a face from the first lineup. The first lineup served to orient them to the test. The methodological reason behind presenting the clone faces was to give sequential participants the opportunity to see which facial features were varying across members in subsequently presented lineups. As the participant progressed through the lineup series, the feature level increased (2, 3, 4, or all 5 features across the faces varied). The goal of the participant was use as few features as possible to correctly distinguish the study face from the other faces. Additionally, the procedure used to present the lineup tests (simultaneous or sequential), and the duration of exposure to the study face (10 s or 30 s) were controlled between participants. In total, participants viewed 6 study faces and were given 6 sets of attendant lineups. A practice trial was given before the experiment began. Participants were verbally instructed throughout.

Participants were seated approximately 40 cm away from a computer screen. After each study face was presented for 10 s or 30 s, a logic puzzle was presented onscreen for 60 s as a distractor task. Participants were told that these "brainteaser problems" were being given to help them pass the time and their answers would not be graded. Thereafter, the lineup tests commenced. A program was created in Visual Basic to display the study faces and lineup tests. The faces were arranged in an array of 2 rows, 3 pictures in each. The faces presented at study were about 18 x 16 cm and the individual pictures presented in the lineup at test were 5.5 x 6.5 cm in size. The position of the target and foils in each lineup was held constant across the lineup series. The onscreen position of the faces within the lineup was maintained in the sequential and simultaneous conditions to control for possible differences in accuracy that might arise from examining faces on different screen locations.

DRAFT

All participants were made aware that from one lineup to the next in a series, the faces would become increasingly different, while the position of the individual faces within the lineup would remain constant. Participants were instructed to identify the face from the earliest lineup in the series as possible while maintaining a high level of accuracy. Once a participant made a decision about a particular face, or lineup, they could not change their mind, or review previously presented faces. If participants positively identified a face, the program would initiate a new study phase, presenting the next face for study. However, if the participant did not identify a face from the lineup, they were given the next lineup in the series. Lineups were presented until a face was identified, or until all five lineups in the series had been presented.

Participants in the sequential condition viewed the lineup faces one at a time, and they were told to indicate for each face whether or not it was one of the faces that they had studied, whereas participants in the simultaneous condition were given instructions to either reject the lineup or identify the study face by position number. Sequential participants could not see the remaining faces in the lineup (if any) once they made a positive identification, nor could they review or change their mind about faces they had rejected previously.

Once the participant expressed an understanding of how the program worked, all participants were given the high criterion instruction that was utilized in Experiment 1.

Measures. Choice rate, accuracy rate, and target selection/choice rates were computed. We also examined for individual participants how they distributed their choices across feature level, and averaged the results across participants within lineup condition. In this manner, we were able to examine whether sequential participants required more evidence (i.e., tended to choose later in the series of lineups) than did simultaneous participants.

DRAFT

Results and Discussion

The choice and accuracy dependent measures were analyzed separately, using a 2 (lineup procedure) x 2 (study face duration of exposure) between subjects ANOVA. The effects of feature level on the target selection/choice data and on the choice distribution data were assessed separately, using a mixed linear model with lineup procedure as a between subjects factor and feature level as the within subjects factor.

The first analysis examined whether participants in the sequential condition required more facial feature evidence before making a selection. Participants in the sequential condition, however, did not request more facial features than simultaneous participants before making a positive identification. Sequential participants tended to choose earlier in the lineup series than those in the simultaneous condition, though the effect was not reliable, $F(3, 166.99) = 1.63$. For participants who chose from a simultaneous lineup, the rate of choosing a face was .14 at the 2nd feature level, .24 in the 3rd, .35 in the 4th, and .26 in the 5th. For sequential choosers, the rates were, in ascending order of feature level, .24, .20, .32, and .23. These results also indicate that the number of test faces viewed in the sequential condition was somewhat lower compared to the simultaneous condition.

In keeping with the hypothesis that sequential lineups induce a more conservative decision standard compared to simultaneous lineups, participants in the sequential condition were less likely to choose a face than those in the simultaneous condition ($M = .43$ and $M = .72$, respectively), $F(1, 96) = 26.99$, $p < .01$, $\eta_p^2 = .22$. Study face exposure duration also significantly impacted the choice rate, such that faces were chosen more often under long ($M = .64$) compared to shorter ($M = .52$) exposure durations, $F(1, 96) = 4.61$, $p < .05$, $\eta_p^2 = .05$. Additionally, study face exposure duration significantly interacted with lineup procedure, which indicated that longer

DRAFT

exposure durations affected choosing in simultaneous lineups (low M = .55 versus high M = .89), but not in sequential lineups (low M = .48 versus high M = .38), $F(1, 96) = 15.41$, $p < .01$, $\eta_p^2 = .14$. These results support the prediction derived from signal detection theory that target discriminability would have a larger effect in simultaneous lineups.

With respect to overall accuracy, lineup procedure, $F(1, 96) = 39.34$, $p < .01$, $\eta_p^2 = .29$ and target exposure duration, $F(1, 96) = 13.17$, $p < .01$, $\eta_p^2 = .12$ had a significant effect. These effects of these variables, however, were qualified by their significant interaction, $F(1, 96) = 17.73$, $p < .01$, $\eta_p^2 = .16$. The interaction obtained for lineup procedure and memory strength occurred because increasing the discriminability of the target via longer exposure durations increased overall accuracy in simultaneous (low M = .22 versus high M = .53) but not sequential (low M = .13 versus high M = .11) lineups. With respect to target selection/choice data, results indicated that for lineup choosers, target selection rates for choosers were higher in the simultaneous compared to the sequential condition (M = .47 versus M = .27), $F(1, 88) = 10.02$, $p < .01$, $\eta_p^2 = .10$, and if the duration of exposure was long rather than comparably shorter (M = .44 versus M = .30), $F(1, 88) = 5.23$, $p < .05$, $\eta_p^2 = .06$. These effects of these variables on target selection/choice rates are qualified by their significant interaction, $F(1, 96) = 4.40$, $p < .05$, $\eta_p^2 = .05$, which indicated that choosers selected the target at a higher rate at longer durations of exposure in the simultaneous condition (M = .60 versus M = .33), but not in the sequential condition (M = .28 versus M = .26).

We conditioned the target selection/choice data on the feature level at which participants chose and compared the results across lineup procedure condition. As shown in Figure 6, target selection/choice rates were not associated with the feature level. Additionally, as shown in

DRAFT

Figure 6, accuracy remained fairly constant in both lineup conditions across feature levels. These results suggest that participants were calibrating the discriminability of the lineup to match the extent to which they could remember the study face.

To summarize the results from Experiment 2, participants who chose early in the series of lineups (and who therefore had less facial feature evidence on which to base their discriminations) were just as accurate as those who chose later in the series of lineups. We propose that these effects occurred because participants were calibrating target discriminability in relation to the degree to which they had learned the study face: Those with weaker memories required additional evidence (i.e., higher target to study face similarity values and lower similarity across members) than those who chose early on.

In keeping with Experiment 1, choice rates were higher in simultaneous compared to sequential lineups, indicating lower criterion placement in the simultaneous condition. Of the participants who chose from the lineup, those in the simultaneous condition were more accurate than those in the sequential condition. These effects were not due to the level of feature evidence requested varying across procedures, as there was no difference based on lineup condition in how participants distributed their choices across feature level. Additionally, lengthening the duration of exposure to the study face increased the accuracy of identifications in the simultaneous but not in the sequential condition.

General Discussion

DRAFT

The results of the experiments demonstrated that correct positive identification rates in simultaneous and sequential lineups are influenced by criterion placement and target discriminability. We will discuss each of these effects in turn.

Previous research has demonstrated that admonishing the witness that the perpetrator may not be in the lineup reduces the false alarm rate (Steblay, 1997) as well as the hit rate (Clark, 2005). Such results are consistent with the view that the decision criterion applied at test is affected by cautionary instructions (Clark, 2005). In the present study, the hit rate for simultaneous and sequential lineups varied depending on whether participants expected that the target was in the lineup. The hit rate did not vary depending on whether the test faces were presented simultaneously or sequentially when participants were instructed that the target was always present in the lineup. In contrast, if participants were admonished that the target may not be present, the decision standard participants implemented varied depending on lineup procedure: Sequential participants set a higher criterion than simultaneous participants, and therefore, they were less likely to choose any face from the lineup, including the target face. These findings suggest that the differences in accuracy that are found across procedures when the target is present are due to differing decision standards, not due to differences in the manner in which information is retrieved from memory.

The criterion shift results also indicated that cautionary instructions not only reduce identifications based on guesses, but sometimes can also discourage witnesses from identifying the target when they do recognize him, a point which was raised by Clark (2005) in his meta-analysis of admonishment effects in lineups. When the criterion was shifted to a more liberal position, target choices increased while foil choices decreased. This latter finding suggests that participants who may have remembered the target in the lineup rejected the lineup rather than

DRAFT

make a positive identification. Additionally, the additional suspect choices that resulted from lowering the decision standard were not simply the result of guessing, otherwise foil choices would have been higher in the low compared to the high admonishment condition. Therefore, at least some of the target misses in the high criterion condition resulted because the target did not surpass threshold, not from memory failure.

We also found that the accuracy rate for choosers was higher if the lineup was administered simultaneously rather than sequentially. Following on Meissner et al. (2005), we propose that these effects occurred because sequential participants could not compare the test faces and make use of perceptual fluency in reaching their decision. We also tested the prediction that sequential witnesses would be less affected by target discriminability than simultaneous witnesses, a prediction that was derived from a signal detection analysis. When target discriminability was operationalized by experimentally controlling the number of features the target shared in common with the study face (Experiment 1), the hit rate rose in the simultaneous and sequential conditions. Though discriminability had less of an effect on hits in the sequential compared to the simultaneous condition, the difference was not statistically significant. When target discriminability was under the control of the participant (Experiment 2), accuracy for choosers was significantly higher for simultaneous compared to sequential participants.

When target discriminability was manipulated via study face exposure duration the rate of accurate identifications was affected in only simultaneous lineups. In the simultaneous condition, the hit rate increased with longer study face exposure durations, whereas no effect of exposure duration was found in the sequential condition. These results are in keeping with the signal detection analysis outlined earlier, which led to the prediction that discriminability would have a larger impact on identifications when the decision standard is set at a relatively lower level.

DRAFT

Meissner et al. (2005) did not find an interaction between lineup procedure and memory strength. Our results may have differed from theirs because of the manner in which target exposure was operationalized; discriminability in the present study was manipulated via an exposure duration manipulation (10 s versus 30 s), whereas they manipulated whether participants studied the face once or twice, with each study period lasting for 3 s, resulting in exposure durations of 3 s versus 6 s. Another key difference between their results and ours is that they reported memory strength effects across target present and target absent lineups combined, where in the current project we analyzed performance as a function of memory strength in target present lineups alone. Future research should follow-up on this result. Ideally, a lineup test procedure should enable investigators to differentiate between witnesses who have stronger memories for the perpetrator. If additional research finds that memory strength has an effect on identifications made in simultaneous but not sequential lineups, then this lends further support to the suggestion made by Doyle et al. (2006) to use simultaneous procedures in practice when other evidence implicating the suspect is high.

In practice, investigators, of course, do not know whether the suspect is definitively guilty. The results from the present project should not be taken to mean that witnesses should be encouraged to guess or provide low confidence positive identifications when presented with a sequential lineup in order to maximize the hit rate. Lowering the decision standard not only increases the probability that the target is correctly identified, but it also increases the probability that an innocent suspect, who looks like the perpetrator, is mistakenly identified. Additionally, a liberal criterion shift induced by encouraging witnesses to identify the suspect even when a witness' confidence is low would produce more false alarms in a simultaneous compared to a sequential lineup procedure (Ebbesen & Flowe, 2002; Meissner et al., 2005).

DRAFT

One possible avenue that might be explored further in laboratory and field research is whether criminal investigators should routinely collect confidence data on all of the test faces when administering a sequential lineup. Previously, many psychologists have viewed confidence as having little to no relationship with accuracy. However, the most recent meta-analysis of the confidence accuracy eyewitness literature (Sporer et al., 1995) indicates that there is a significant positive association between accuracy and decision confidence for *choosers*, or people who chose a face from a lineup. Additionally, although it was a smaller effect than the one obtained for choosers, the Sporer et al. 1995 meta-analysis also indicated a significant positive association between confidence and accuracy for *nonchoosers* (i.e., people who do not identify anyone from the lineup). Since confidence does appear to have some relationship with accuracy, it might be useful to routinely obtain confidence ratings from actual eyewitnesses. Psychologists typically measure confidence using an interval rating scale, whereas in practice, police usually obtain a nominal confidence rating (Behrman & Richards, 2005). To be sure, the legal system is interested in obtaining definitive answers from an eyewitness ("Is the perpetrator present in the lineup, yes or no?" "Are you sure that's him?") as a vehicle for providing probative evidence to the court about criminal identity. However, if a high criterion witness recognizes the target in a sequential lineup, but does not positively identify him because the decision threshold is not exceeded, an opportunity to further investigate the suspect may be gained if decision confidence based on an interval scale is also obtained. Presumably in such circumstances, the witness would assign a lower confidence rating to the target face compared to the foil faces, which would be of value to investigators. Additionally, in cases with multiple witnesses, obtaining decision confidence may be beneficial. Confidence ratings in these circumstances may increase the informative value of the combined testimony, say, if one witness makes a positive identification

DRAFT

and another does not. Other avenues that could be explored by additional research in the lab and in the field to increase guilty suspect identifications include: utilizing simultaneous lineups in conjunction with strong admonishment instructions; altering the simultaneous procedure such that witnesses do not know in advance how many faces they will see; and/or requiring simultaneous witnesses to provide a yes/no decision to each face, as suggested by Zimmerman, Malpass and MacLin (2008). Recent research they have conducted indicates that these methods may hold promise for increasing the reliability of eyewitness identifications.

There are limitations in the current set of studies that need to be addressed. The effects we report may be exclusive to target present lineups. Decision processes may very well differ when the target is absent from the lineup (see Clark, Howell, & Davey, 2008). For instance, participants may engage in secondary search strategies when there is not a face in the lineup that matches the target in memory. It has been demonstrated that extended searches, such as taking multiple "laps" through the faces in a lineup, is associated with increased false alarms (MacLin & Phelan, 2007). Additionally, there were procedural differences between the sequential lineup procedure we used and the sequential lineup procedure that has been recommended by practice guidelines (TWGEE, 1999). Namely, sequential participants in our study were necessarily aware of the number of faces in the lineup because they were shown multiple lineups. In practice, eyewitnesses should not be aware of the number of test faces that will be shown so as to prevent them from selecting faces at the end of sequence simply because they know that no other faces will be displayed (TWGEE, 1999). To prevent this occurrence, sequential lineups should be *backloaded*, which means that witnesses should be led to think that they are going to see more faces then they actually will see. The reduction in guessing as a consequence of backloading would be equivalent to setting a higher criterion in signal detection terms. Therefore,

DRAFT

backloading the lineups in the present study would have decreased the hit rate in both simultaneous and in sequential lineups. Zimmerman, Malpass and MacLin (2008) found that backloading influenced responding in both simultaneous and in sequential lineups; however, it is not clear yet as to whether these results apply exclusively to target absent lineups, because they found significant effects of backloading in target absent but not target present lineups. Further research is clearly needed.

Another potential limitation concerns the type of face stimuli that were employed. One may argue that the decision processes observed in the present study are unique to composite faces and may not obtain with photographs of actual faces. We believe that there is ample evidence to doubt that this is the case. First, composite face stimuli and photographs of faces may activate the same regions of the brain. Ellis and Young (1988) found they were unable to train a child with prosopagnosia (which is a disorder that is associated with brain damage, which behaviorally manifests as the inability to identify familiar faces) to recognize or draw either real or schematic drawings of faces. These results suggest that schematic faces and real faces are processed in the same regions of the brain. Second, composite drawings of faces have revealed insights about face processing (e.g., Leder & Bruce, 2000; Tanaka & Farah, 1993; Tanaka & Sengco, 1997), and the effects have not differed depending on whether composite faces or photographs of faces were used as stimuli. Third, sequential superiority effects (Flowe & Ebbesen, 2007) and inversion effects (Howarth & Flowe, 2009) have been obtained with FACES. Taken together, these findings indicate that the composite drawing face stimuli that we used in the present study are processed in a similar manner as photographs of actual faces. Still, additional research is warranted with face stimuli of increasing complexity, defined by distinguishing features, hairstyles, and facial hair. Currently, there are no studies that have compared the effects of these

DRAFT

factors in relation to performance in simultaneous and sequential lineups. These factors may affect accuracy differentially depending on the type of lineup procedure that is utilized. It will be interesting indeed to see how such factors contribute to distinctiveness and face processing in lineup tests (see Gronlund, 2005).

Finally, additional work is needed to examine the relationship between target-culprit similarity and identification accuracy. In the present study, the similarity of the lineup target compared to the study face was manipulated by systematically varying that number of features on the test face that matched the study face. As a consequence, participants may have relied on local features in recognizing previously studied faces to a greater extent than would otherwise be the case. Specifically, though a person can be recognized on the basis of having matching facial features (e.g., Brunas, Young, & Ellis, 1990; Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985), configural-based processing plays a large role in face recognition (Maurer, Le Grand, & Mondloch, 2002). However, there is no empirical evidence to suggest that the relative contributions of feature- versus configural-based processing varies across lineup procedure. Therefore, the feature-based method by which target-"culprit" similarity was manipulated in the present study seems adequate at this time for examining the effect of target discriminability on target selections in simultaneous and sequential lineups.

In sum, the research reported in the present paper indicates that lineup procedure and admonishment instructions in conjunction with target discriminability have interesting interactive effects on recognition accuracy in target present lineups. The results further indicated that there is some room to increase the hit rate in target present lineups by attending to these factors. Future research should continue to investigate methods for so doing.

DRAFT

References

Behrman, B. W., & Richards, R. E. (2005). Suspect/foil identification in actual crimes and the laboratory: A reality monitoring analysis. *Law and Human Behavior, 29*, 279-301.

Brunas, J., Young, A. W., & Ellis, A. W. (1990). Repetition priming from incomplete faces: Evidence for part to whole completion. *British Journal of Psychology, 81*, 43-56.

Carlson, C., Gronlund, S., & Clark, S. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 14*(2), 118-128.

Clark, S.E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior, 29*, 395–424.

Clark, S. E., Howell, R. T., & Davie, S. L. (2008). Regularities in Eyewitness Identification. *Law and Human Behavior, 32*, 187-218.

Cnaan, A., Laird, N. M. and Slasor, P. (1997). Using the general linear model to analyze unbalanced repeated measures and longitudinal data. *Statistics in Medicine 16*, 2349-2380.

Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73*(2), 281-290.

Doyle, J. M.., Penrod, S., Kovera, M. B., & Dysart, J. (2006). The Street, The Lab, The Courtroom, The Meeting Room. *Public Interest Law Reporter, 11*(2), 13-46.

Ebbesen, E. B., & Flowe, H. D. (2002). Simultaneous v. Sequential Lineups: What Do We Really Know? Retrieved November 13, 2008, from http://www-psy.ucsd.edu/%7eeebbesen/SimSeq.htm.

DRAFT

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*, 431-439.

Ellis, H. D. & Young, A. W. (1988). Training in face-processing skills for a child with acquired prosopagnosia. *Developmental Neuropsychology, 4*(4), 283-294.

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*(1), 33-52.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). *Remembering and knowing*. New York, NY, US: Oxford University Press. .

Gronlund, S. D. (2005). Sequential lineup advantage: Contributions of distinctiveness and recollection. *Applied Cognitive Psychology, 19*(1), 23-37. .

Hanley, J. A., Negassa, A., deB. Edwardes, M. D., & Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: An orientation *American Journal of Epidemiology, 157*, 364-375.

Howarth, D. E. & Flowe, H. D. (2009). Composite and inversion effects examined in simultaneous and sequential lineups: Is there a difference between procedures in configural/holistic processing? Manuscript in preparation.

Leder H., & Bruce V. (2000). When inverted faces are recognized: the role of configural information in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*, 513-536.

Levi, A. M. (1998). Protecting innocent defendants, nailing the guilty: A modified sequential lineup. *Applied Cognitive Psychology, 12*(3), 265-275.

Lindsay, R. C., Lea, J. A., Nosworthy, G. J., & Fulford, J. A. (1991). Biased lineups: Sequential

DRAFT

presentation reduces the problem. *Journal of Applied Psychology, 76*(6), 796-802.

Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Leach, A., & Bertrand, M. (2009). Sequential lineup presentation: Patterns and policy. *Legal and Criminological Psychology, 14*(1), 13-24.

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

MacLin, O. H., & Phelan, C. M. (2007). PC_Eyewitness: Evaluating the New Jersey Method. *Behavioral Research Methods, 39*(2), 242-247.

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

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

Malpass, R. S., Tredoux, C. G. & McQuiston-Surrett, D. (2009). Public Policy and Sequential Lineups. *Legal and Criminological Psychology, 14*(1), 1-12

Maurer, D., Le Grand, R., & Mondloch, C.J. (2002). The many faces of configural processing. *Trends Cognitive Science, 6*(6), 255–260.

McQuiston-Surrett, D.E., Malpass, R.S., & Tredoux, C.G. (2006). Sequential vs. simultaneous lineups: A review of methods, data, and theory. *Psychology, Public Policy and Law. 12*(2), 137-169.

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.

DRAFT

Mickes, L., Wixted, J. T., Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review, 14*(5), 858-865.

Patterson, K.E & Baddeley, A.D (1977) When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3*(4), 406 - 417

Ratcliff, R., Sheu, C. F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518-535.

Read, J. D., Vokey, J. R., & Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 870-882.

Sporer, S. L. (1994). Decision times and eyewitness identification accuracy in simultaneous and sequential lineups. In David Frank Ross (Ed.), *Adult eyewitness testimony: Current trends and developments.* (pp. 300-327): Cambridge University Press, New York, NY, US.Sporer, S., Penrod, S., Read, D., & Cutler, B.L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin , 118*, 315–327.

Steblay, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law & Human Behavior, 21*(3), 283-297.

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

DRAFT

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.

Tanaka, J. W., & Farah, M. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46*, 225-245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition, 25*, 583-592.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* [Booklet]. Washington, DC: United States Department of Justice, Office of Justice Programs, National Institute of Justice.

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

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

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

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*(1), 52-176.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language, 46*, 441-517.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin, 133*(5), 800-832.

DRAFT

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A.W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*, 737–746.

Zimmerman, L. A., Malpass, R. S., & MacLin, O. H. (2009). Unconfounding the simultaneous v. sequential lineup comparison. *Manuscript under review*.

DRAFT

Figure Captions

Figure 1. Application of signal detection theory to simultaneous (M) and sequential (Q) lineup identification when the target is present.

Figure 2. Receiver operator characteristics (ROCs) for lineups in which target discriminability and criterion placement vary in sequential and simultaneous lineups.

Figure 3. Illustration of stimuli that were used in Experiments 1 and 2. Target appears in position 5 of the lineup (i.e., in the middle of the bottom row).

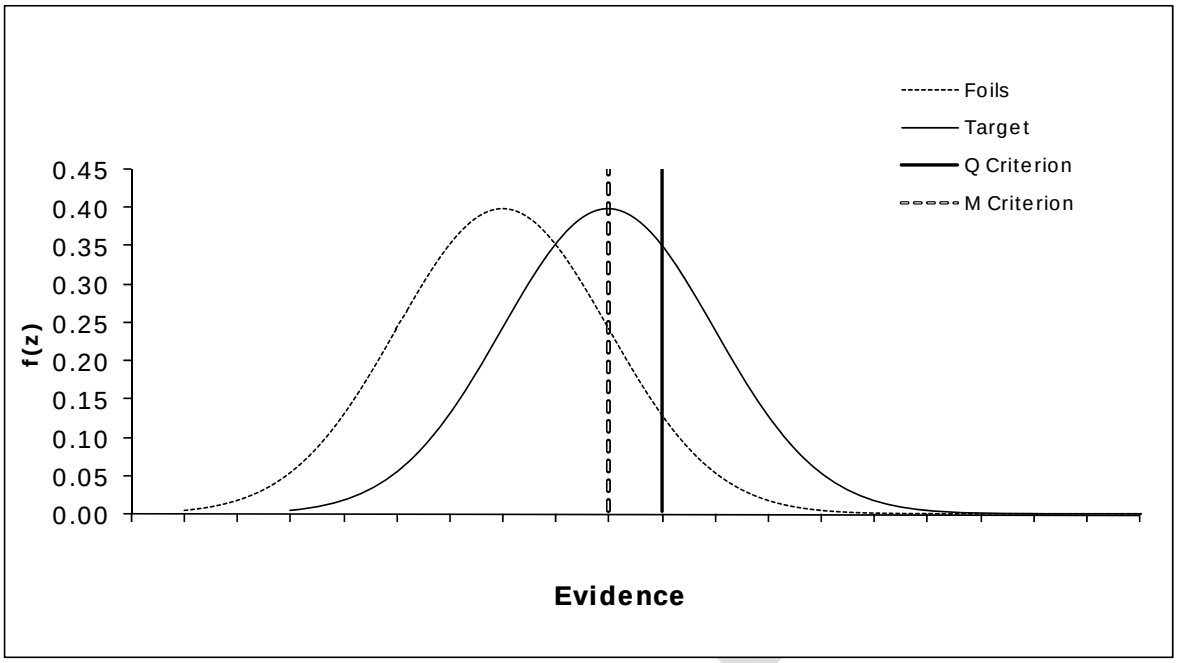
Figure 4. Mean (±1 *SEM*) response data (x-axis) by feature level (y-axis) for foil and target choices in sequential and simultaneous lineups under low versus high criterion instructions in Experiment 1.

Figure 5. ROCs derived from confidence ratings made in positive identifications from sequential and simultaneous lineups in Experiment 1.

Figure 6. Mean (±1 *SEM*) target ID/choice data across feature level for simultaneous and sequential lineups in Experiment 1.

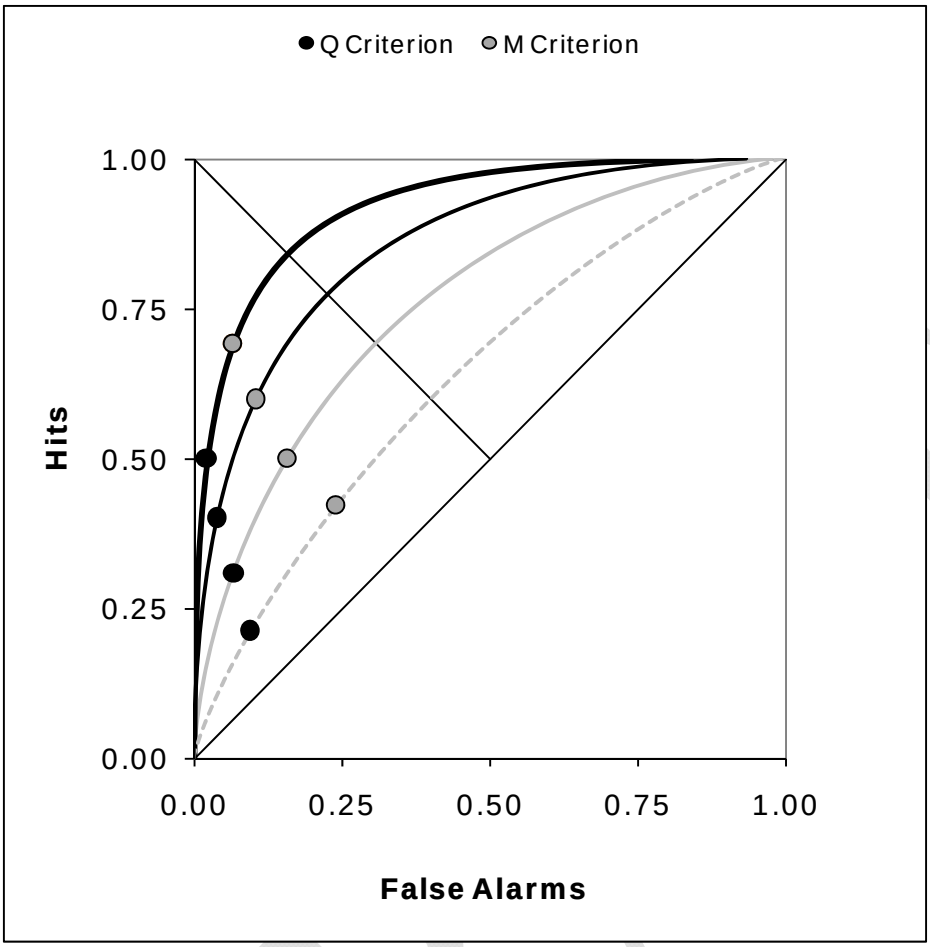
DRAFT

Figure 1



DRAFT

Figure 2



DRAFT

Figure 3

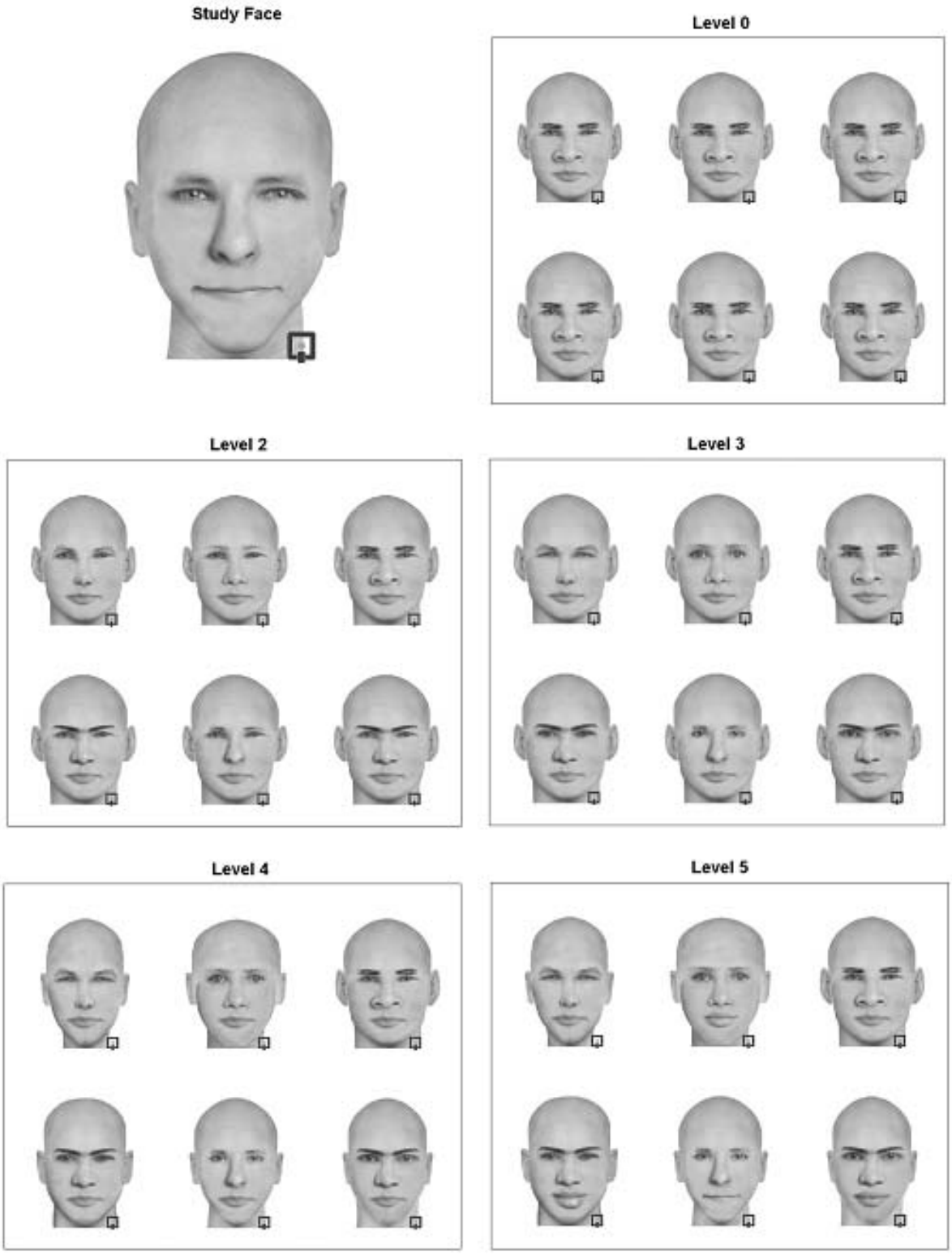
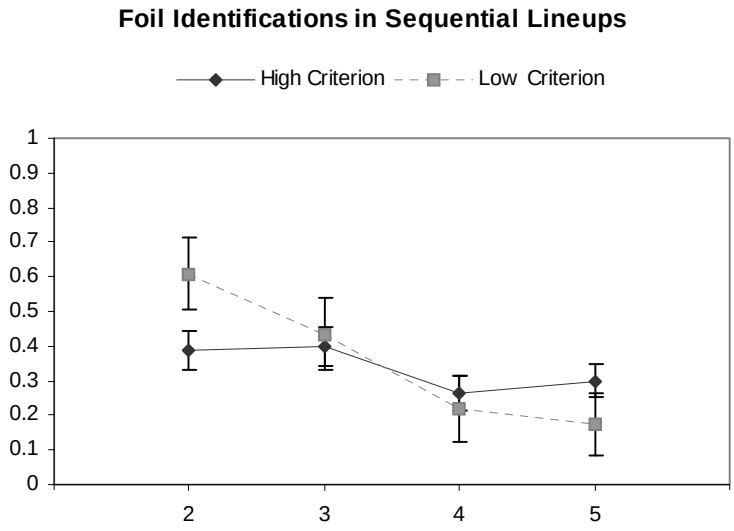
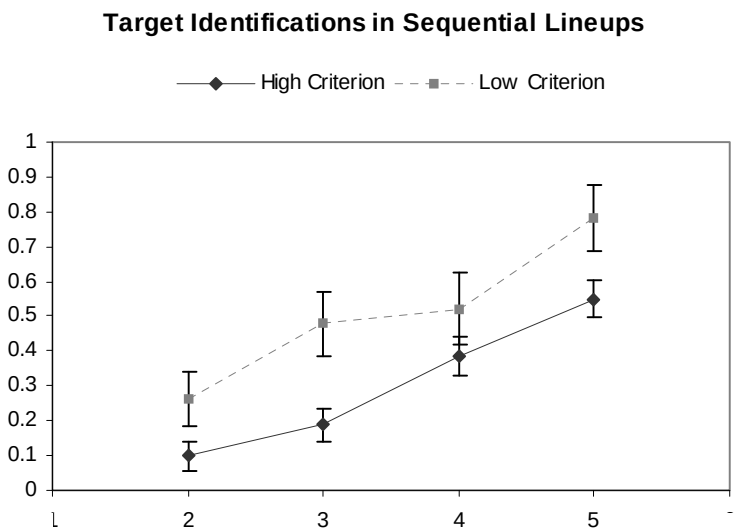
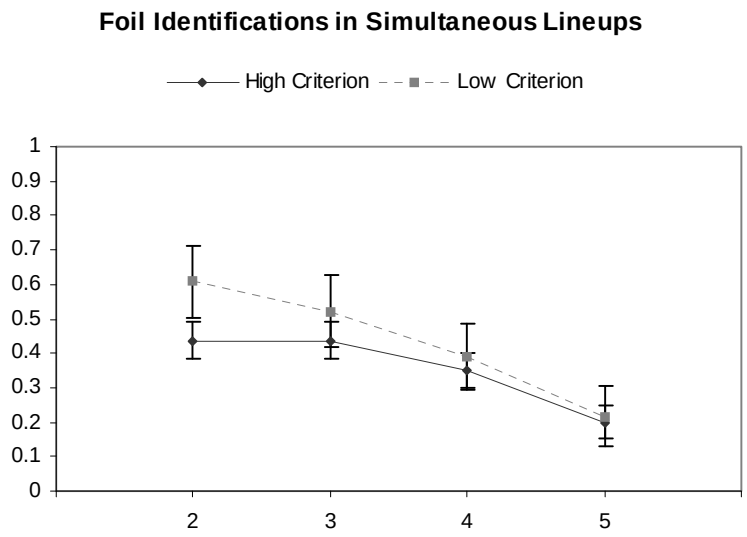


Figure 4


Foil Identifications in Sequential Lineups
High Criterion
Low Criterion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5


Target Identifications in Sequential Lineups
High Criterion
Low Criterion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5


Foil Identifications in Simultaneous Lineups
High Criterion
Low Criterion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5

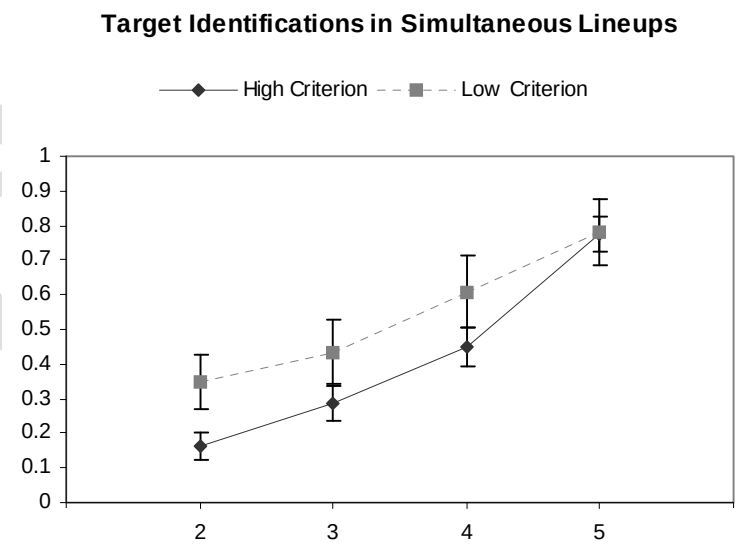
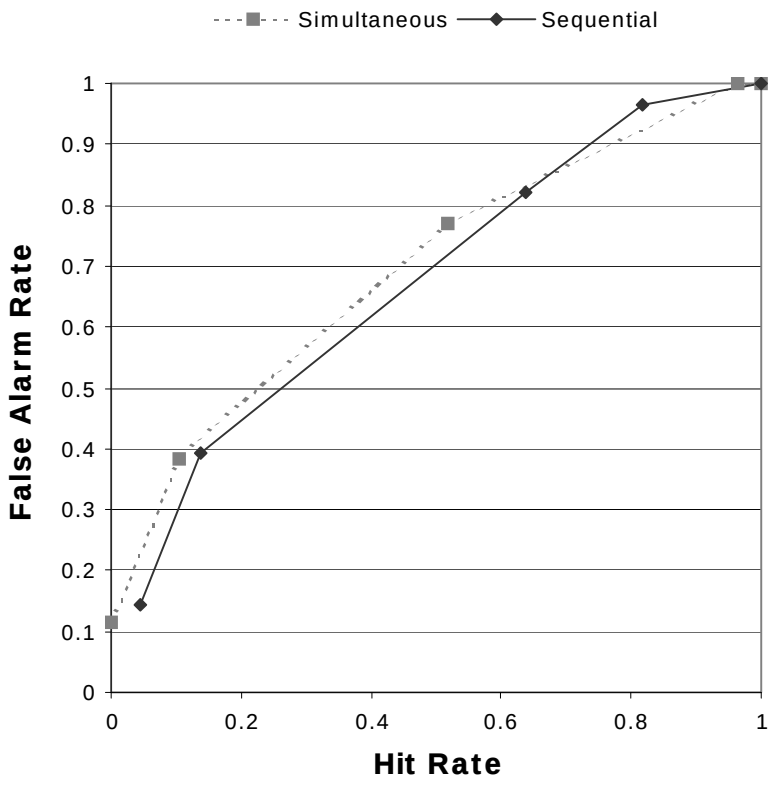
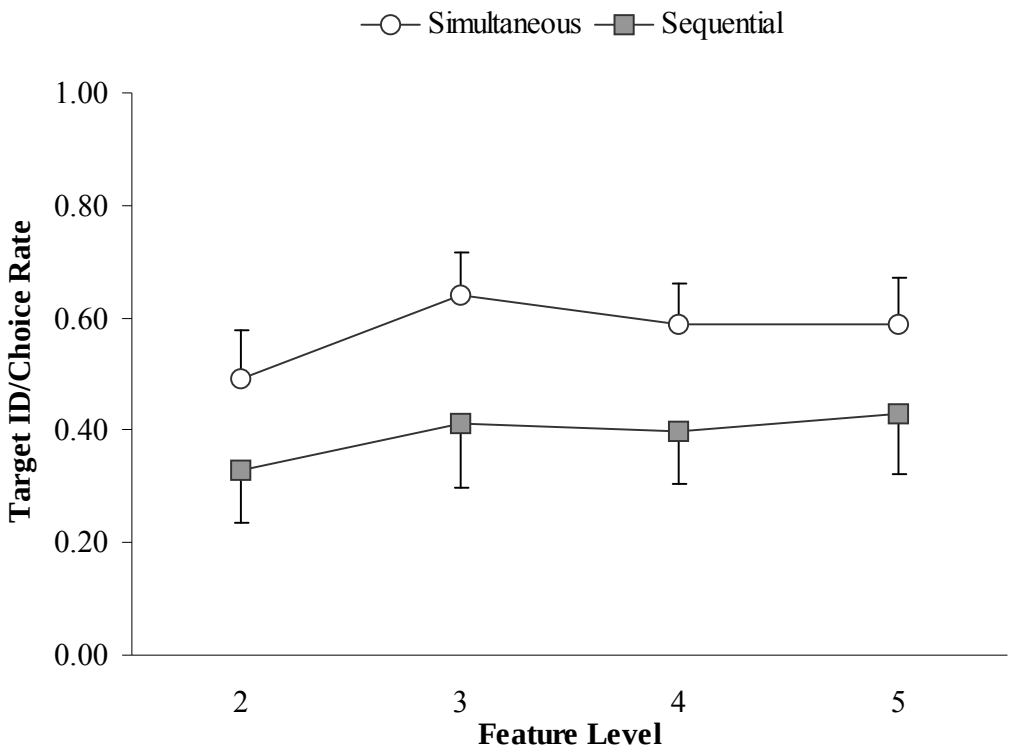

Target Identifications in Simultaneous Lineups
High Criterion
Low Criterion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5

Figure 5



DRAFT

Figure 6



DRAFT