

The Effect of Lineup Member Similarity on Recognition Accuracy in Simultaneous and Sequential Lineups

Heather D. Flowe · Ebbe B. Ebbesen

Published online: 4 October 2006

© American Psychology-Law Society/Division 41 of the American Psychological Association 2006

Abstract Two experiments investigated whether remembering is affected by the similarity of the study face relative to the alternatives in a lineup. In simultaneous and sequential lineups, choice rates and false alarms were larger in low compared to high similarity lineups, indicating criterion placement was affected by lineup similarity structure (Experiment 1). In Experiment 2, foil choices and similarity ranking data for target present lineups were compared to responses made when the target was removed from the lineup (only the 5 foils were presented). The results indicated that although foils were selected more often in target-removed lineups in the simultaneous compared to the sequential condition, responses shifted from the target to one of the foils at equal rates across lineup procedures.

Keywords Witnesses · Legal testimony · Judgment · Memory · Face perception

Much psychological research over the last 25 years has demonstrated the conditions under which eyewitnesses might be prone to making identification errors. Based on what we have learned from these studies, psychologists have called on the legal system to make procedural changes aimed at reducing the rate of mistaken identification (Wells et al., 1998). In particular, psychologists have begun to recommend the use of the sequential lineup over the traditional simultaneous lineup as a means of reducing false identifications (e.g., Lindsay, Lea, Nosworthy, & Fulford, 1991; Wells et al., 1998).

The sequential procedure differs from a simultaneous lineup in a number of respects. In a simultaneous lineup, witnesses are shown at once an array of faces, which includes the suspect along with persons known by the police to be innocent (foils). In a sequential lineup, the faces are viewed one at a time, and for each face the witness makes a yes/no decision. The next picture is displayed if the photo is rejected by the witness, and once a photo has been rejected, the witness is not allowed to see it again. The sequential lineup test continues until a face is positively identified as the culprit. Across laboratory studies, it has been shown that by switching

H. D. Flowe (✉) · E. B. Ebbesen
Department of Psychology, University of California,
San Diego, 9500 Gilman Drive, La Jolla, CA 92093
e-mail: hflowe@ucsd.edu

to sequential lineups, identifications of both “guilty” and “innocent” suspects are reduced, though false identifications of suspects are reduced to a larger extent (see Steblay, Dysart, Fulero, & Lindsay, 2001) for a meta-analytic comparison of simultaneous and sequential lineups).

One explanation that has been proposed to account for the differences in accuracy observed between the two procedures is the relative-absolute judgment model (Wells et al., 1998). According to this model, positive identifications are higher in simultaneous lineups because witnesses can see all of the alternatives when making their identification decision, and as a result, are biased toward making a positive identification. Simultaneous witnesses will likely make a positive identification because at least one of the faces should be a relatively better match to the culprit in memory than the other faces in the array. Reliance on a relative judgment strategy can account for the typical finding that false alarms to the innocent suspect, who presumably should look relatively more like the culprit than any of the other lineup members, occur with greater frequency in simultaneous lineups. Relative judgments are thought to figure less prominently in identification decisions made from sequential lineups because direct visual comparisons of lineup members cannot be made. Instead, sequential witnesses must rely on an absolute judgment strategy, which entails matching the given face against the culprit in memory, and evaluating whether the match evidence exceeds the criterion threshold for a positive identification. Wells et al. (1998) note that although previously seen faces could theoretically influence whether a given face is identified from a sequential lineup, witnesses will rely more on an absolute strategy because they cannot be sure that one of the faces that they have not seen yet will be an even better match to the culprit than the one that they are currently evaluating.

Little is known, however, about the degree to which the similarity structure of the lineup faces (i.e., the similarity of the foils with respect to one another and the culprit) presented earlier in a sequential lineup affects judgments made to later faces. At first thought, this may seem inconsequential in predicting suspect identifications. If the relative comparison process is impeded by presenting faces one at a time, then whether a given face is selected should not be influenced by the other members in the lineup. On the other hand, the similarity of previously seen faces relative to the culprit could influence judgment if comparisons were made between the currently presented face and the previously seen lineup faces that are in working memory. In this paper, we set out to test whether lineup similarity structure has a comparable effect on remembering in simultaneous and sequential lineups. Specifically, will the rate of choosing the face that is the best match to the target be influenced by the similarity of the other faces to the target in both lineup procedures? Two experiments were conducted to address this question. An overview of our predictions and the relevant literature are presented next.

Effects of similarity on criterion placement in identification tasks

Research has shown that criterion placement is affected by characteristics of the lineup task, such as nominal size (Xu & Bellezza, 2001), admonishment instructions that warn witnesses that the perpetrator may not be in the lineup (Clark, 2005; Steblay et al., 2001), and clothing bias (Lindsay, Wallbridge, & Drennan, 1987). We propose that the similarity structure of the lineup also influences criterion placement. In particular, we predicted in the current project that foils who are low in similarity to the culprit lead witnesses to adopt a more liberal decision criterion, thereby increasing the probability that a suspect who looks similar to the culprit will be identified. In this section we review literature that supports this proposition.

It seems clear that in both simultaneous and sequential lineups, participants set a decision criterion for making a positive identification. For instance, in both laboratory (Steblay et al., 2001) and archival studies (Behrman & Davey, 2001), simultaneous witnesses reject the lineup without identifying anyone about 50% of the time. If simultaneous witnesses were simply

picking the relatively best face, they would almost always select someone. Therefore, any model of eyewitness identification from lineups, whether simultaneous or sequential, will have to include a criterion parameter.

Ebbesen and Flowe (Ebbesen & Flowe, 2002) cast simultaneous and sequential lineups within a signal detection framework, predicting that sequential participants utilize a more conservative decision criterion than simultaneous participants, as across studies, participants are less likely to pick any face from the lineup when administered sequentially. Meissner, Tredoux, Parker, and MacLin (2005) confirmed this prediction, finding that while discrimination accuracy did not vary between the two procedures, a conservative criterion shift occurred when faces were presented sequentially rather than simultaneously. Moreover, witnesses might adopt an even more conservative criterion compared to simultaneous and sequential lineups when presented with a showup identification task, which involves presenting a single face to an eyewitness who then makes a yes/no culprit identity decision. Meta-analytic work finds that showup choice rates are lower when compared to simultaneous or sequential lineups in both target present and absent conditions (Stebly, Dysart, Fulero, & Lindsay, 2003). The lower choice rates in showups compared to lineups is consistent with the idea that the similarity structure of a lineup created by the addition of foils might affect criterion placement.

Additional support for the proposition that lineup similarity affects criterion placement is found in the verbal learning literature. Recent studies have demonstrated that the plausibility of the distractors presented on a final recognition test affects criterion placement (Benjamin, 2005; Benjamin & Bawa, 2004). These studies show that increasing the degree of semantic overlap between the distractors and the previously learned material increases participants' decision criterion, as both hits and false alarms on the recognition test are reduced. These findings suggest that participants evaluate mnemonic evidence from the test items to gauge how easily they can discriminate the targets from the lures and then set their decision criterion accordingly. Applying this finding to simultaneous and sequential lineups leads to the prediction that as the similarity of the foils to the culprit increases, the odds that the suspect (whether guilty or innocent) will be identified will decrease in both procedures. Of course, simultaneous witnesses are immediately able to gauge task difficulty because all of the test faces are in full view, whereas sequential participants must wait for the lineup to unfold. Nevertheless, if sequential witnesses make meta-comparisons while responding to a given face on the test, then the similarity structure should influence their decisions as well.

In keeping with this idea, in the eyewitness literature, (Clark & Davey, 2005) found evidence suggesting that sequential witnesses are sensitive to the similarity characteristics of the lineup. They employed the removal without replacement procedure (Wells, 1993) to test whether sequential participants make relative judgments from lineups. The removal without replacement procedure involves comparing the number of choices made to the most popular foil in a lineup when the target is present, to the number of choices made to that same foil when the target is removed from the lineup (and not replaced by any other face). When the target was removed from the lineup, simultaneous and sequential witnesses shifted, to the same degree, their choice to the foil who looked the most similar to the culprit. They proposed that this indicates that a common decision process underlies remembering faces in simultaneous and sequential lineups.

Clark and Davey (2005) further showed that when the similarity of the foils in the lineup was reduced relative to the culprit, the most similar-looking foil (relative to the culprit) was chosen at a higher rate if he appeared in the order later (position 4) rather than earlier (position 2). They invoked a within lineup criterion shift explanation to account for the findings. Specifically, when the next best alternative was presented early in the lineup, participants withheld making a choice in order to find out whether a better option would be presented later. Only low similarity foils appeared thereafter, and consequently participants rejected the lineup. In contrast, if low

similarity foils appeared first and the next best alternative was in position 4, participants lowered their decision criterion and chose him. In the current study, we expand on these findings by investigating whether surrounding the target (or the look-a-like) with foils that are high or low in similarity influences the rate at which the target is chosen. If the similarity of the foils influences criterion placement in simultaneous and sequential lineups, the target and the look-a-like should be chosen at a higher rate in both procedures if surrounded by low rather than high similarity foils.

To our knowledge, only Lindsay et al. (1991) has systematically varied lineup similarity along with lineup presentation procedure to examine their joint effect on accuracy. Participants were presented with lineups that had foils that were all similar or dissimilar to a confederate in a staged-crime experiment (experiment 3). Though foil choices were increased overall in the similar relative to the dissimilar condition, only lineup procedure alone affected false identifications. Since other studies have demonstrated low foil-culprit similarity can increase false identifications in simultaneous lineups (e.g., Wells & Lindsay, 1980), in the present study, we used multiple study faces, look-a-likes, and lineups to increase the generality of the similarity manipulations over faces. In so doing, we were able to examine similarity effects across several different lineups.

Overview of experiments

In the first experiment, participants studied a set of faces and then attempted to identify them from a lineup, which was presented either simultaneously or sequentially. The lineups contained two dimensions of variation: similarity of the foils to the study face, and the similarity of the suspect in the lineup to the study face. Showup tests were also utilized, and results from these were compared with the lineup identification outcomes to estimate the degree to which the foils enhanced or detracted from remembering. In addition, composite drawings of faces were utilized as face stimuli so as to tightly control similarity, since other methods, such as creating a photo lineup based on the suspect's description or a match to the suspect's photo, have produced somewhat inconsistent results across studies (see Tunnicliff & Clark, 2000). In a subsequent section we will discuss the possible limitations that might arise as a result of using composite faces rather than photographs of faces.

In the second experiment reported, the removal without replacement procedure was employed to determine whether simultaneous and sequential participants shifted their choice to a foil rather than rejecting the lineup. Additionally, a group of subjects rank ordered the similarity of the study face to the lineup foils to determine which foil was the next best alternative. These rankings were compared to identification outcomes in simultaneous and sequential lineups to see if they predicted foil choices in both procedures.

Method

Participants

A total of 294 people participated in exchange for course credit. One participant was removed from the final analysis due to experimenter error. The final sample was 86% female, and in total, 58% of the sample was Asian, 34% Caucasian, 6% Hispanic, and 2% self-identified as other.

Design

The identification procedure (simultaneous, sequential, and showup) and lineup feature similarity structure (random and matched) variables were fully crossed to create 6 conditions to

which participants were randomly assigned. Each participant was given 12 study faces and accompanying identification tests. This procedure might be likened to a real world case in which there is more than one perpetrator and witnesses are given an identification test for each suspect (e.g., *People vs. Cobb*). For half of the identification tests, the identical study face was present. For the other half, a highly similar looking face was substituted in place of the study face to create a target absent condition. Lastly, we nested within the matched lineup similarity structure condition the facial feature used to match the foils in the lineup to the study face (eyes, mouth, or face).¹

Stimulus materials

Face stimuli were created using FACES, a composite drawing software program used by law enforcement. Combined over all types of facial features (eyes, nose, mouth, face, eyebrows, jaw and head shape, hairstyle and facial hair), the program contains a databank of over 3,850 facial components that can be used to construct a composite drawing. Within each facial feature category, the user can select from a range of morphological sizes (i.e., the default size of a chosen feature, such as the nose, can be enlarged) and types (i.e., the shape of a particular feature, such as bulging or almond shaped eyes). The program can also generate at random a complete composite face. To create the study faces and accompanying lineups in the present study, a pool of 1,000 faces was created by having the program generate faces at random. All distinguishing features, facial hair, and head hair were removed from each face.

Simultaneous and sequential lineups

The lineups for the random similarity structure condition were formed by randomly selecting without replacement 6 faces from the initial pool of 1,000. The only restriction on the faces selected for the random lineups was that the eyes had to be the same color (light or dark) and no facial feature could match across members. From these 6 faces, the to-be-identified study face was randomly designated. Using this procedure, 12 study faces and 12 accompanying lineups were created.

For the matched similarity structure condition, the same procedure was employed to select 6 lineup members, though rather than the features across the lineup faces varying at random, the lineup faces were matched on one feature. Three versions of each lineup were created. Each version was based on the particular feature (eyes, mouth, and face shape) that the foils had in common with the target face (see Fig. 1 for an example of the lineups used for one of the study faces). In each feature match condition (eyes, mouth, and face shape), the foils matched the study face on only one feature; the remaining features for each of the foils were the same as the original version selected from the stimulus pool. In the eye matched condition, for instance, each

¹ In a pilot study, we presented participants with lineups in which the target face was a mirror image of the face that had been studied. Our purpose in so doing was to ensure that participants were making their identifications based on their recognition of facial features rather than on some other more general pattern matching strategy. However, we learned in post study interviews that participants were reluctant to make an identification if they did not find a face that precisely matched the study face. For instance, if a slight fold appeared under the right eye only, mirroring the face resulted in the fold appearing under the left eye. In such cases, we found that when participants recognized a face as a mirror image of the study face, many would not identify it because it was not a perfect match. In fact, a small number asked spontaneously during test whether they should identify a mirror image. Thus, in the final study reported here, test pictures were not mirrored from the original. In addition, the size of the lineup test pictures was reduced to about a third of the size of the study pictures to make identifications based on pattern matching alone more difficult.

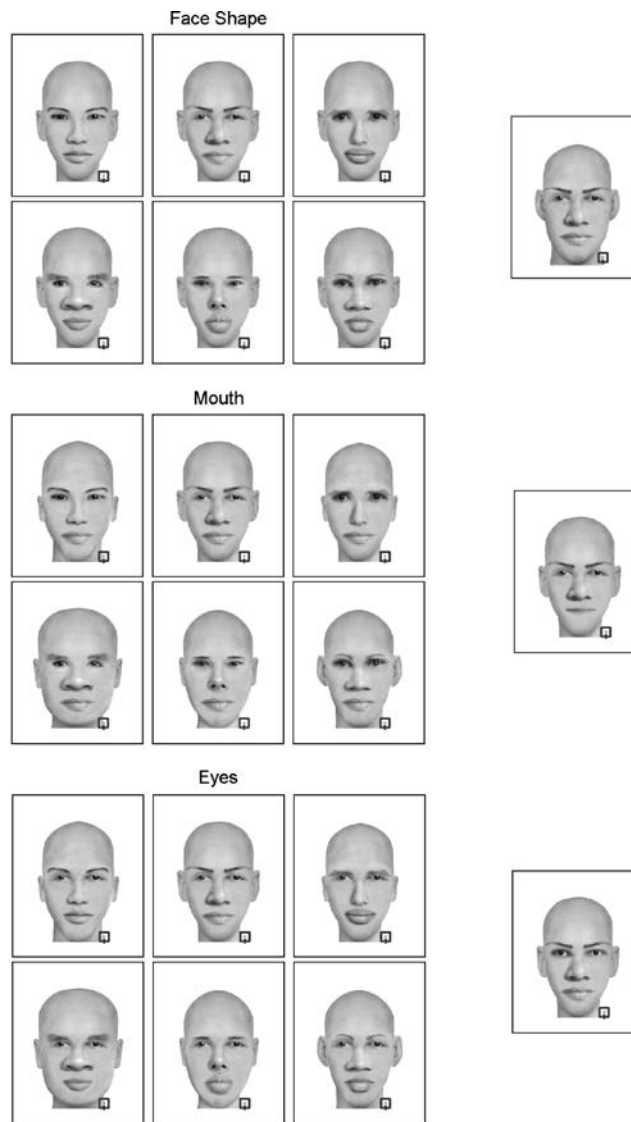


Fig. 1 Example of lineup stimuli for a study face (*row 1, center*) in the matched condition for each feature level, which is indicated at the top of the lineup. The look-a-like that took the place of the study face is located to the right of each lineup

of the foil faces in the lineup was altered from the original such that the eyes were now identical in shape and size to the target face. Note that in this condition the eyes were now identical across all foils in the lineup. Using this procedure, 12 study faces and three accompanying lineups for each were created. Participants in the matched condition were randomly assigned 12 lineups at test, with the constraint that they view one lineup for each study face and a total of four lineups for each of the 3 feature match conditions.

In both the random and matched conditions, the suspect (i.e., study face or look-a-like) appeared early (positions 1–3) in the array for half of the study faces and late (positions 4–6) for the other half. The order in which the foils appeared in the lineups was randomly determined. The position of the target and foils in each lineup was held constant across subjects. Furthermore, for the three versions of matched lineups made for each study face, position of the lineup members was maintained. If the lineup test was simultaneous, the pictures were presented in 2 rows of 3 pictures. The faces presented at study were about 18 cm × 16 cm and the individual pictures presented in the lineup at test were 5.5 cm × 6.5 cm in size. A number corresponding to the position of the face in the lineup (1–6) was placed beneath each picture for purposes of the identification task. For the simultaneous lineups, numbering was from left to right, starting on the top row.

To manipulate the similarity of the lineup target to the study face, a highly similar version of the study face (hereafter referred to as the “look-a-like”) was created for both the random and matched conditions. This new version of the study face was created by deleting one feature from the original study face and substituting it for another. The single feature deleted was the eyes, face shape, or mouth. The feature substituted in place was randomly determined. See Fig. 1 for examples of the matched condition.

To create the target absent lineups used in the random similarity structure condition, study faces were randomly assigned to have one of the three types (eyes, face shape, and mouth) of feature deletions. As a result, study face and the feature used to create the accompanying target absent version of the lineup were confounded in the random condition.

In the matched condition, the feature deleted was counterbalanced across the 12 study faces. The altered feature was identical to the one that had been used to match the foils. For example, if the foils had the same eyes as the study face, then the suspect look-a-like would have different eyes than the study face (and hence the foils in the lineup).

For the random and matched lineups containing the look-a-like, the study face was removed from the lineup and the similar looking target was placed in the same position. Each participant was given six target absent lineups, and of these, two were of each feature substitution type.

Showups

For both the random and matched conditions, the faces designated as the suspects for the simultaneous and sequential lineups were presented alone as a showup test. Half of the tests presented the study face, whereas the other half presented the look-a-like.

Similarity manipulation check

As a check on our lineup similarity manipulations, another group of participants ($n = 120$) made pairwise similarity ratings for all of the faces in each lineup (15 comparisons possible per lineup). Participants were randomly assigned to evaluate 12 lineups; half evaluated matched lineups and the other half random lineups. The face pairs were presented in random order, one pair at a time, with the faces in the pair presented simultaneously on a computer screen while participants gave their ratings. Ratings were based on a similarity scale (that ranged from 0 to 100, with 0 being “not at all similar” and 100 being “completely the same”). The experimenter walked participants through three practice trials, encouraging them to base their ratings on the facial features of the lineup members. Thereafter, participants made the ratings on their own and at their own pace.

Another group of participants ($n = 120$) rated the similarity of the feature substituted look-a-likes with respect to the study faces. They also made a total of 180 pairwise similarity comparisons, 36 of which were comparisons of the study faces and the look-a-likes used in the present study ($3 \text{ look-a-likes} \times 12 \text{ study faces}$). The remaining pairs of faces ($n = 144$) that they evaluated were drawn from lineups that were used in an earlier pilot study of this project.²

Finally, we also ran a simultaneous mock witness task to determine whether the similarity structure of the lineups enabled participants to readily recognize the suspect in absence of any memory for the suspect. Toward this end, another group of 72 participants was randomly assigned to evaluate either random or matched lineups, half of which were target present. The experimenter explained that in a previous study, participants had studied a set of faces and then had their recognition memory for the faces tested with lineups. In the present study, their task was to try to determine which face, if any, was the target face in each of the lineups. They were instructed to not identify anyone if none of the faces seemed to stand apart from the others, because the study face was sometimes not present in the lineup in the previous study.

Procedure

At the start of the study, participants in all three identification test conditions (simultaneous, sequential, and showup) were simply asked to pay close attention to the faces that they were about to see, that they would be asked questions about them later. Study faces were then presented in random order for 10 s each followed by a 5-min retention interval, during which time participants completed a crossword puzzle. Participants were verbally instructed that the study face may or may not be present.

Participants were correctly informed that the order in which the lineups (or showups) were presented did not necessarily match the order in which the faces had been studied. In addition, for the simultaneous and sequential conditions, they were told that there would be only one person, if any, in the lineup that was a previously studied face. In other words, there would never be more than one study face presented in a given lineup.

Those in the sequential lineup condition were informed that for each lineup, they would only see one picture at a time and make a yes/no decision regarding whether a given face had been presented at study. They were informed that once they made a decision, the decision would stand. They would not be allowed to view previously seen pictures, nor would be they be allowed to change any of their answers. In addition, if they positively identified a face, the test would terminate—even if there were faces in the lineup that remained to be seen—and the next identification test would be given. Some investigators have allowed sequential participants to see the remaining faces in the lineup after making a positive identification to an earlier face (e.g., Clark & Davey, 2005). One reason for so doing is to determine whether participants who identify a foil early in the sequence can recognize the study face if it appears later in the sequence. The identification test, however, is effectively over after the first identification. If a real world witness, for instance, identified multiple faces, he or she would be discredited. In addition, we were concerned that showing all of the faces after an earlier one was identified might affect criterion placement in our study because participants were given multiple lineup tests rather than just one. Specifically, if participants learned that they could identify a face and still see the

² The standard that this group of participants used to evaluate the similarity of the study face and look-a-likes did not seem to differ from the standard used by the other group of participants who evaluated the full lineups from the present study. The average similarity of the 144 faces that we are disregarding did not differ from the average similarity of the faces from the lineups used in the present study.

others in the lineup, we worried that they might place their criterion in a more liberal position, and thereby pick multiple faces in every lineup. Thus, we did not allow participants to see the remaining faces once an identification was made.

Another feature that distinguishes our sequential lineups from those used in practice is that our participants had the opportunity to learn that only 6 faces would be shown in each lineup. In practice, sequential witnesses should not know the number of faces that they will see, so as to avoid an increased tendency to pick members at the end of the sequence simply because participants know that no other pictures will be displayed. This issue will be considered in the results wherein we present an analysis of choice rates by position number to determine whether our procedures influenced participants to choose more often in the later part of the lineup sequence.

For all identifications, testing was not timed. The experimenter waited until the participant made a selection before moving on to the next test. If a participant asked whether they made a correct decision, the experimenter would indicate that she did not know, or that she could not tell him until the end of the study.

Dependent measures

Overall accuracy was calculated separately for the study face and look-a-like conditions for the simultaneous ($n = 94$), sequential ($n = 93$), and showup ($n = 134$) participants. In the study face condition, selecting the target was coded as an accurate response, while selecting a foil or rejecting the lineup was coded as an incorrect response. For the look-a-like condition, rejecting the lineup was coded as an accurate response, while selecting a foil or the similar looking target was counted as an incorrect response.

Two more dependent measures were calculated: (1) Choice rate, or the mean proportion of responses in which any face was positively identified and (2) target choice/ID, or the mean proportion of responses in which the target face was chosen given the participant positively identified any face in the lineup.

Results

Preliminary analyses

First we checked whether the targets utilized in the random compared to the matched condition affected the results. No significant effect of similarity condition on accuracy (random $M = .64$ versus match $M = .66$) in showups was obtained. Therefore, any differences between random and matched lineups reported in subsequent analyses cannot be attributed to pre-existing differences between the conditions in the memorability of the study faces. In addition, the particular feature used to create the matched lineups did not differentially affect accuracy depending upon lineup procedure. Therefore, data from the matched condition were collapsed across feature in subsequent analyses.

Next we examined whether the pattern of responses varied over the course of the 12 trials. Since participants were given 12 faces to remember followed by 12 lineup tests, this raises the possibility that the demands of the task may have led to cognitive overload, thereby inducing sequential participants to adopt a relative decision strategy. We tested this possibility by examining choice and accuracy rates as a function of trial number. The probability of selecting the suspect did not vary as participants progressed through the series of lineups, as the correlation between suspect choices and trial number was not significant (simultaneous $r = -.28$, sequential

$r = -.30, p$'s > 0.33 , two-tailed). Though not significant, the negative trend suggests that cognitive load was increasing over the course of the lineup tests. This, however, did not seem to lead to relative responding. If participants were adopting a relative strategy as a result of fatigue, then the rate of choosing any face should have increased over the 12 trials. Performance, however, seemed to remain constant over the 12 trials, as the correlation between trial number and choice rates was not significant in sequential ($r = .09, p = .77$, two-tailed) or in simultaneous ($r = .02, p = .93$, two-tailed) lineups.

We also compared the rate at which sequential participants chose faces in a given lineup position (early, positions 1 and 2, middle, positions 3 and 4, and late, positions 5 and 6) on the first and last trial. The distribution of choices on the first compared to the last trial were not different (Trial 1: early $M = .22$, middle $M = .21$, and late $M = .07$; Trial 12: early $M = .24$, middle $M = .18$, and late $M = .08$). In addition, lineup rejection rates on the first and last trial did not differ (Trial 1: $M = .47$ and Trial 12: $M = .51$). These data suggest that sequential participants maintained a similar decision strategy over the course of the lineup tests.

Finally, we assessed the similarity ratings to determine whether our similarity manipulations were instantiated as intended. As expected, the average similarity across all 6 members of the lineups was significantly higher in the matched (39.63, $SD = 9.63$) compared to the random (30.72, $SD = 4.02$) condition, $t(1,82) = 3.14, p < .05$, two-tailed, indicating that the manipulation worked as intended. Additionally, feature substitutions made to the target face significantly affected the similarity structure of the lineup. The average similarity of the foils relative to the study face (38.71, $SD = 7.84$) was higher than the average similarity of the foils relative to the look-a-like (33.81, $SD = 8.16$) in the matched condition, $t(1,70) = 2.57, p < .05$, two-tailed. Given that we used a 101 point scale, the ratings may seem to indicate that the similarity of even the matched lineups was pretty low. However, the average similarity of the single feature removed look-a-likes relative to their original study face was 70.48 (average similarity by feature condition: eye $M = 68.29, SD = 8.83$; face $M = 74.63, SD = 9.62$; mouth $M = 68.50, SD = 8.96$). The level of these average ratings suggests that the raters were somewhat reluctant to use the upper range of the similarity scale.

With regard to the simultaneous mock witness task, we analyzed whether suspect choices varied as a function of lineup structure and/or target condition. The results showed that mock participants did not select the suspect at a rate greater than chance (.17) in any of the conditions (Match: target present $M = .13$, target absent: $M = .17$; Random: target present $M = .16$, target absent: $M = .16$). These results indicate that in the absence of memory for the study faces, simultaneous witnesses could not discern which member was the focus of the lineup. Consequently, if simultaneous witnesses in the experiment proper choose the suspect at a rate greater than sequential witnesses, the result cannot be attributed to simply an unfair advantage in ability to guess correctly in simultaneous lineups because all of the choices were laid out.

Taken together, the results of the similarity manipulation checks suggested that we were in a good position to test how the similarity structure of the lineup affected identification responses. We turn to these results next.

Feature substitution effects in showups

We wondered if the feature substitutions made to the study faces would significantly decrease their familiarity when presented alone in a showup. Participants were more apt to correctly reject the look-a-like ($M = .72$) than they were to correctly identify the study face ($M = .59$), $F(1,132) = 16.59, p < .01, \eta_p^2 = .11$. This result also implies that

Table 1 Identification Responses by Lineup Construction, Identification Procedure, and Target Level for Experiment 1

Lineup construction and ID procedure	Target	Suspect	Foil	Reject
Matched				
Sequential	Identical	0.37 (.03)	0.27 (.03)	0.36 (.03)
	Similar	0.13 (.03)	0.30 (.03)	0.57 (.03)
Simultaneous	Identical	0.45 (.03)	0.26 (.03)	0.29 (.03)
	Similar	0.24 (.03)	0.32 (.03)	0.44 (.03)
Random				
Sequential	Identical	0.33 (.05)	0.30 (.04)	0.37 (.04)
	Similar	0.24 (.05)	0.36 (.04)	0.40 (.05)
Simultaneous	Identical	0.52 (.05)	0.23 (.04)	0.25 (.04)
	Similar	0.34 (.05)	0.27 (.04)	0.39 (.05)

Note. Data shown are the mean rate (*SE*) of suspect picks, foil picks, and lineup rejections.

showup participants were biased toward rejecting rather than positively identifying the test faces.

Accuracy in simultaneous and sequential lineups

Adding foils to the lineup seemed to reduce correct responses, as mean accuracy was significantly lower in sequential (.41) and simultaneous lineups (.46) compared to showups (.65), $F(1,317) = 82.49, p < .01, \eta_p^2 = .21$. In showups, positive identifications of the study face were significantly higher, while identifications of the look-a-like were significantly lower than in simultaneous or sequential lineups (all p 's $< .05$). These results indicate that identifications were affected by the foils, regardless of whether they were presented simultaneously or sequentially.

As the means presented in Table 1 illustrate, participants were less likely to reject a simultaneous compared to sequential lineup ($M = .33$ versus $M = .43$, respectively), $F(1,183) = 6.85, p < .01, \eta_p^2 = .04$. In addition, a higher rate of choosing occurred in random compared to matched lineups, (simultaneous: random $M = .68$ and match $M = .63$; sequential: random $M = .62$ and match $M = .54$), $F(1,183) = 4.91, p < .05, \eta_p^2 = .03$. The difference in choice rates indicates that lineup similarity structure had an effect on decision criterion placement. Next we examined how the structure of the lineup affected accuracy in each of the lineup procedure conditions.

If similarity affects decision criterion placement, suspect choices should increase as the foils become less similar to the study face. Participants were more likely to select the study face in a random compared to matched simultaneous lineup, though the effect was marginally significant, $F(1,92) = 3.35, p = .07, \eta_p^2 = .04$. The rate of selecting the study face in sequential lineups was unaffected by similarity condition, $p = .47$. With respect to the look-a-like, the face was picked at a significantly higher rate in a random compared to matched lineups in the simultaneous condition, $F(1,92) = 6.65, p < .05, \eta_p^2 = .07$, as well as in the sequential condition, $F(1,91) = 9.79, p < .01, \eta_p^2 = .10$.

ID target/choice data were examined to determine under what conditions targets drew more choices than foils. The similarity rating data suggest that foil choices should be higher in matched compared to random lineups. Compared to the study face present condition, foils drew choices away from the suspect to a greater extent when the look-a-like was in the lineup. The rate of choosing the suspect rather than a foil was significantly higher in the study face condition ($M = .60$)

compared to the look-a-like condition ($M = .39$), $F(1,173) = 39.32$, $p < .01$. With respect to the target absent lineups, foils drew choices away from the look-a-like more often in the matched compared to random condition in both simultaneous and sequential lineups, though the differences were not significant: Choosers did not select the look-a-like significantly more often in random compared to matched lineups in the simultaneous condition (random $M = .48$; match $M = .39$) or in the sequential condition (random $M = .37$; match $M = .30$), all p 's $> .27$.

Position effects by similarity condition

The sequential lineup data were further examined to determine whether the position of the suspect influenced accuracy rates. Specifically, suspects appearing late rather than early might be missed more often because an earlier presented foil who looks similar to the target is chosen instead. Suspect choice data were entered into a mixed ANOVA, with target condition (study face or look-a-like) and suspect position (early or late) as the within subjects factors, and lineup construction (matched or random) as a between subjects factor. A significant two-way interaction was obtained for position and lineup construction, $F(1,69) = 5.84$, $p < .05$, $\eta_p^2 = .08$. Participants were more likely to pick the suspect if presented early ($M = .48$) in the sequence rather than late ($M = .36$) in the matched condition ($p < .05$), while the opposite held true in the random condition (early $M = .24$ versus late $M = .30$), though the difference was not significant in the latter condition ($p = .24$). The only other significant effect in the analysis was an interaction between target condition and target position, $F(1,69) = 5.98$, $p < .05$, $\eta_p^2 = .08$. The study face was chosen at a higher rate in early compared to later positions ($M = .47$ and $M = .36$, respectively), while the look-a-like was chosen equally often in late compared to early positions ($M = .24$ and $M = .29$). Finally, the position effects found for sequential lineups were not due to the memorability of the particular targets that were placed in early compared to later positions, as no significant position effects were obtained in simultaneous lineups (early $M = .43$ versus late $M = .48$), $p = .27$.

Summary

Experiment 1 investigated whether lineup similarity structure affected the distribution of choices in lineups. In particular, we tested the hypothesis that surrounding the suspect with low compared to high similarity foils will result in lower criterion placement. Similarity effects were examined in both simultaneous and sequential lineup procedures because different decision processes are thought to underlie identifications in the two procedures. A relative decision strategy, which is used to characterize simultaneous identifications, should increase suspect choices as the similarity of the foils relative to the perpetrator is decreased because witnesses can readily ascertain the similarity characteristics of the lineup and set their identification criterion accordingly. In sequential lineups, the structure of the lineup should have little consequence for the rate at which the target is identified if judgments are based primarily on an absolute strategy. In such cases, only the similarity of the suspect to the perpetrator should influence positive identifications of the suspect.

Experiment 1 found that accuracy in showups was higher compared to simultaneous and sequential lineups, indicating that the addition of foils affected remembering. Moreover, lineup similarity structure affected accuracy, regardless of whether the members were presented simultaneously or sequentially. Specifically, when a look-a-like was in the lineup, participants

were more apt to select him if the foils were lower in similarity to the study face. When the study face was present, however, the structure of the lineup had a marginal effect on accuracy in simultaneous lineups, and no effect on accuracy in sequential lineups.

That similarity affected false alarms to the look-a-like more than hits to the study face is not necessarily problematic for our hypothesis. Given that the study faces were more memorable than the feature substituted look-a-likes, criterion adjustments should have had a greater effect in the target absent compared to the target present condition (see Ebbesen & Flowe, 2002). In addition, in real world lineups, it is unlikely that a guilty suspect in a lineup will appear exactly how he or she looked while committing the crime. In the target present lineups we used, the targets were identical in appearance to the faces that participants studied. Further research is needed to examine whether the similarity effects we obtained occur in target present lineups in which the “guilty” suspect in a lineup differs in appearance compared to when participants studied him. In such cases, similarity effects in target present lineups might be larger than what we observed, because familiarity of the target will be reduced by changing his appearance during the lineup task.

Additionally, position effects were observed in sequential lineups in keeping with findings from (Clark & Davey, 2005). In the matched condition, the foils were higher in similarity to the study face than in the random condition. Consequently, some sequential participants in the matched condition selected one of the feature matched foils presented early in the sequence, thereby missing the later appearing target. In random lineups, the suspect was chosen at a higher rate if presented later compared to earlier in the sequence, though the difference was not significant.

In Experiment 2, we investigated the effects of lineup member similarity on lineup choices in closer detail. The removal without replacement procedure (Wells, 1993) was employed, using the lineups and study faces from Experiment 1. The distribution of responses that were made to lineups when the target was removed and the 5 foils were presented alone (hereafter referred to as the “target-removed” condition) was compared to the responses that were made in Experiment 1 when the target was present (hereafter referred to as the “target-present” condition). In particular, we examined whether the most widely chosen foil in the target-present lineups was selected at an even higher rate in simultaneous and sequential target-removed lineups.

In Experiment 2, we also investigated whether rank orderings as a measure of lineup member similarity predicted foil choices in target-present and target-removed simultaneous and sequential lineups. Toward this end, an independent group of participant raters rank-ordered the similarity of the study face to the lineup foils. If the most highly ranked foil corresponded to the most often identified foil in the target-present lineups in both lineup procedures, then this suggests faces are processed in a similar manner in sequential and simultaneous lineups. Such regularity would not be expected if different match to memory evidence was used during the test in sequential compared to simultaneous lineups.

The ranking method was used over other similarity measures for two reasons. First, rank ordering required participants to make relative comparisons. Experiment 1 demonstrated that lineup identifications are influenced not only by the similarity of each face to the perpetrator, but also by the similarity of the other members. As such, it might be important to capture this process in measures of lineup similarity by having raters make assessments of foil-perpetrator similarity with all of the lineup faces in view. Second, the ranking method was chosen to check whether the most widely chosen foil was the lineup member ranked as the most similar to the perpetrator. For instance, it is possible that the most widely identified foils in Experiment 1 were chosen more often than the other members simply because they were the most distinctive faces in the lineup, not because they were the most similar to the suspect.

Experiment 2

Participants

A total of 208 undergraduates participated for course credit. Among these, 72 participated in the identification portion of the study, while 136 rank-ordered the faces in the lineups.

Design

The protocol for the recognition part of the study was similar to Experiment 1, with identification procedure (simultaneous and sequential) and lineup similarity (random and matched) fully crossed to form four experimental conditions to which participants were randomly assigned.

With respect to the similarity ranking portion of the study, another group of participants was provided with a target face and told to rank order the members in the corresponding lineup from most to least similar. Half of these participants ranked random lineups and the other half ranked the matched lineups. In addition, participants in the matched condition were randomly assigned to make rankings for 12 lineups, with the constraint that they view one lineup for each study face and a total of four lineups for each of the three feature match conditions. Those in the random condition viewed the 12 random lineups that were presented in Experiment 1. In both the random and matched conditions, the lineups were presented to each participant in a random order.

Procedure

The protocol used in the identification portion of the study was identical to that employed in the first experiment, except that participants were tested with target-removed lineups.

In the ranking portion of the study, participants were asked to rank order the lineup members from 1 to 5, with 1 designated as “most similar” and 5 “least similar” to the target face. Tied rankings were not allowed. Participants were allowed to look at the target face as often as they wished while completing the ranking task. A total of 8 participants were removed from the analysis because instructions were not followed (e.g., assigned tied ranks, did not rank all of the faces).³

Results

The results are divided into two parts. In the first, foil and lineup rejection responses are compared in simultaneous and sequential lineups. In the second part of the results, the similarity ranking data are compared to foil choices in the target-present (TP) and target-removed (TR) simultaneous and sequential lineups.

³ Another group of participants studied the 12 faces and then made the similarity rankings for the 12 lineups based on their memory for the study faces. The memory rankings for each lineup were significantly correlated with similarity rankings that were made based on direct visual comparisons.

Table 2 Mean Rate (*SE*) of Foil Identifications by Condition for Simultaneous and Sequential Lineups in Experiment 2

ID procedure and lineup condition	Popular foil	Other foil	Reject
Sequential			
Full	0.18 (.02)	0.15 (.02)	0.42 (.02)
Removed	0.20 (.02)	0.32 (.02)	0.47 (.02)
Simultaneous			
Full	0.14 (.02)	0.15 (.02)	0.34 (.02)
Removed	0.22 (.02)	0.47 (.02)	0.34 (.02)

Note. Data shown are the mean rate (*SE*) of most popular foil picks, other foil picks, and lineup rejections.

Choices in removed lineups by lineup procedure

Data were analyzed at the level of study face (collapsed across the 3 feature conditions in the matched lineups) to ensure that one of the foils would emerge as the clear favorite (hereafter this foil is referred to as the “most popular” foil). Foil choice rate data were entered into a mixed ANOVA, with lineup condition (TP versus TR), and foil condition (most popular foil versus other foil choice) as repeated measures and lineup similarity and identification procedure as the between groups factors (see Table 2 for the results). A total of 2 simultaneous and 6 sequential lineups were not included in the analysis because no foil emerged as a clear favorite in the TP condition (i.e., responses were distributed equally, or nearly so, across the foils). Only significant effects that emerged from the analysis are reported.⁴

Removing the target from the lineup significantly increased foil choices, $F(1,36) = 148.57$, $p < .01$, $\eta_p^2 = .80$. A significant interaction was obtained, however, for lineup condition and identification procedure, indicating that foils were chosen at a higher rate when the target was removed in simultaneous (TP $M = .14$ versus TR $M = .34$) compared to sequential lineups (TP $M = .16$ versus TR $M = .26$), $F(1,36) = 15.82$, $p < .01$, $\eta_p^2 = .31$. Since the hit rate was larger in simultaneous lineups in Experiment 1, there was greater opportunity for choices to shift in simultaneous compared to sequential lineups when the target was removed. The degree to which choices shifted to a foil can be captured by forming a ratio of the difference in TP and TR foil choice rates to the target identification rate (Wells, 2006). Expressed in this way, the target to foil shift (on average across lineups) was 106% (.39/.38, $SE = .10$) in the simultaneous condition and 84% (.20/.26, $SE = .13$) in the sequential condition, a difference that was not statistically significant, $t(38) = -1.24$, $p = .11$, one-tailed. Considered together, these results indicate that although foils were chosen at a higher rate in the simultaneous compared to the sequential condition when the target was removed, the proportion of hits that shifted to fillers did not differ by lineup procedure.⁵

The next set of results indicates to which foil participants directed their choice when the target was removed. Other foils were chosen at a significantly higher rate than the most popular foil, $F(1,36) = 19.30$, $p < .01$, $\eta_p^2 = .27$. This main effect is qualified by the significant interaction

⁴ Based on this coding, the most popular foil in TP lineups corresponded with the most popular foil in TR lineups 57% of the time in the sequential condition, and 46% of the time in the simultaneous condition. In TP lineups, 26% of the time the most popular foil corresponded in the simultaneous and sequential conditions. Larger agreement between the two lineup procedures in foil choices was found in TR lineups, as 56% of the time the most popular foil in simultaneous lineups was also the most popular foil in sequential lineups.

⁵ We would like to thank the anonymous reviewer who suggested this analytical approach.

that was obtained for foil condition and lineup condition, which indicated that the rate of selecting the most popular foil from the TR lineups did not increase as much as the rate of selecting one of the other foils (most popular foil: full $M = .16$ and removed $M = .21$; other foil choices: full $M = .15$ and removed $M = .39$), $F(1,36) = 65.32$, $p < .01$, $\eta_p^2 = .64$. Though the interaction shows that the increase was greater for other foil choices, follow-up analyses indicated that the most popular foil was chosen at a significantly higher rate ($p < .01$) when the target was removed. Additionally, foil condition significantly interacted with lineup procedure, $F(1,36) = 4.96$, $p < .05$, $\eta_p^2 = .12$. The interaction emerged because other foils were chosen at a higher rate in simultaneous ($M = .31$) compared to sequential lineups ($M = .23$), while the rate of choosing the most popular foil did not vary depending on lineup procedure ($M = .18$ simultaneous and $M = .19$ sequential).

Ranking data compared to foil identifications

To determine which lineup foil was the most similar to the study face, the ranking data within each lineup (random $n = 12$, matched $n = 36$) were averaged across raters. The lineup member with the highest average rank in each lineup was deemed as the most similar foil with respect to the study face. Lineups in which more than one foil was tied for the highest rank were included in the analyses. If the most widely chosen foil in the full (or removed) lineup was one of the highest ranked foils in the tie, it was coded as a case in which the ranking corresponded with the foil choice data. Foil choice data were collapsed across target present and absent lineups, as the pattern of correspondence between the rankings and foil identifications was the same, regardless of whether the target was present or not.⁶

The most popular foil identified from the TP lineups received the highest ranking in 36% of the simultaneous lineups and in 28% of the sequential lineups; these rates were not reliably different $\chi^2(1) = 0.98$, $p = .32$. The correspondence between the rankings and the most widely chosen foil in TR lineups was higher in simultaneous than in sequential lineups, though the difference was not reliable. The memory rankings and foil identification correspondence rates were 48% in simultaneous lineups and 35% in sequential lineups $\chi^2(1) = 2.65$, $p = .11$.

Summary

Simultaneous participants chose foils at a higher rate in target-removed lineups; however, the degree to which choices were shifted from the target to a foil did not differ across lineup procedures. In both simultaneous and sequential target-removed lineups, other foils were identified more often than the most popular target-present lineup foil. This finding may have emerged because the foils in each lineup were evenly matched in their similarity to the target. As such, responses were distributed somewhat evenly across the fillers rather than being directed at any one filler in particular. In addition, similar levels of agreement across lineup procedure were obtained between the ranking data and filler choices. This suggests that similar memory to match evidence was being used in making identifications from simultaneous and sequential lineups.

⁶ We repeated the analyses reported in this section excluding the lineups in which there was a tie for the highest ranked foil and the pattern of results was the same.

General discussion

The results of Experiment 1 indicated that criterion placement can be affected by the similarity structure of the lineup. In both simultaneous and sequential lineups, the look-a-like was selected at a higher rate in lineups in which the foils were lower in similarity to the study face. The rate of choosing any face was also higher if the foils were lower in similarity to the study face, indicating lower criterion placement. Moreover, the results suggest that sequential witnesses make meta-comparisons of some sort during the identification task. Such a process might entail making relative comparisons (in working memory) similar to the manner outlined by (Wells et al., 1998). Meissner et al. (2005) suggested that another plausible mechanism involved in the lineup recognition process might be perceptual fluency, a term which refers to the ease of perceptual processing when familiarity is used as a determinant of recognition (Jacoby & Dallas, 1981). Relative to the low similarity foils, perceptual processing of the suspect might be easier than when high similarity foils appear in the lineup. If relative ease of processing is used as a cue to establish the probability that a face is the perpetrator, then positive identifications of the suspect are expected to increase.

Experiment 1 also found that the similarity of the foils differentially affected accuracy in sequential lineups depending on where the “suspect” was placed in the lineup. Specifically, when the foils were high compared to low in their similarity to the study face, sequential participants tended to select a foil presented earlier in the sequence, effectively missing suspects placed in later positions. Though the rate of selecting the suspect was higher if he appeared later rather than earlier in a low similarity lineup, the difference was not significant. The pattern of findings is consistent with (Clark & Davey, 2005). They found that the foil most similar to the culprit was chosen more frequently in early compared to later positions when high similarity foils were in the lineup (experiment 1). In contrast, the most similar foil was chosen more frequently in later compared to earlier positions when lower similarity foils were in the lineup (experiment 2). Taken together, these findings indicate that the position of the suspect affects accuracy differently, depending on the similarity structure of the lineup. A topic for further study is whether the similarity of the first face shown in the sequence affects criterion placement for subsequently viewed faces, or whether sequential participants adjust their criterion only after seeing several faces that are dissimilar to the culprit. In addition, degree of similarity might be an important factor warranting further study, as the effects observed in the present study could be larger when the similarity of the foils to the target is even higher.

The results of this project can be compared to prior studies undertaken to examine how suspect choices are affected by matching foils to a witness’ description of the perpetrator as opposed to matching the foils to the suspect’s photo. Lusk and Wells (1993) were the first to hypothesize that identification accuracy might be reduced in target present lineups if the foils are chosen based on their match to the culprit’s photo rather than a witnesses description of the culprit. In addition, if the foils in the lineup match the suspect’s photo, and the suspect is innocent and closer in resemblance to the culprit than the other members, then the suspect will be the member chosen most often. In such cases, the suspect will be relatively more similar to the culprit than the other members.

However, results across studies comparing identification outcomes in description matched and suspect matched lineups have been somewhat mixed and may be attributable in part to differences across studies in the way in which similarity was controlled across foils and suspects (see Tunnicliff & Clark, 2000) for a review). If description matched lineups on average produce a low level of foil-culprit similarity and suspect matched lineups produce a higher level of similarity, then the results from the current study lead to the prediction that false alarms will be higher in description-matched lineups. In description matched lineups, witness’ criteria for

making a positive identification will be lower than in suspect-matched lineups. Moreover, given that the culprit is well-learned, identifications from target present lineups will not be as strongly affected as identifications from target absent lineups by the similarity structure of the lineup (see Ebbesen & Flowe, 2002). These predictions should hold, of course, only when the innocent suspect resembles the actual culprit. Luus and Wells (1993) maintain that the innocent suspect will look no more like the actual culprit than do any of the foils if the lineup is description matched. Further research is needed to establish the degree to which the look-a-like must match the culprit (in terms of both the witness' memory and objectively similar physically) in order for the similarity of the foils to have an effect on accuracy. The look-a-likes used in the current project were highly similar to the study face, and therefore, our results only apply to such lineups.

The results also suggest that test faces are analyzed in a similar manner in simultaneous and sequential lineups. In Experiment 2, similarity rankings of the foils to the study face predicted which of the foils would be chosen most often in both lineup conditions. In addition, consistent with (Clark & Davey, 2005), we found both simultaneous and sequential participants shifted their choice to a foil when the study face was removed from the lineup. Foil choices were greater for simultaneous participants, owing to the fact that criterion placement in simultaneous lineups is lower compared to sequential procedures. However, the degree to which responses shifted from the target to one of the foils did not differ based on lineup procedure. These results suggest that information is extracted from faces in a similar manner during simultaneous and sequential lineup tests.

Given the nature of our stimuli and procedures, a number of possible limitations need to be addressed. First, the results of the current project are limited to lineups in which the foils are evenly matched to the perpetrator. The degree to which the foils are uniformly similar to each other has been found to influence identifications in other work (Laughery, Jensen, & Wogalter, 1988). Second, the use of composite drawings to control similarity raises the concern that the results may not apply to photographs of actual faces. However, findings regarding the impact of foil similarity on accuracy in lineups with composite drawings have been replicated using photographs of faces (Wogalter, Marwitz, & Leonard, 1992). In addition, ongoing research in our lab using photographs of faces selected from sex offender databases has replicated the finding that in both simultaneous and sequential lineup procedures, lower similarity lineups result in more false alarms to look-a-likes compared to high similarity lineups.

Other possible limitations of the current work center around the procedures employed, namely, the fact that we had participants study 12 study faces and take 12 lineup tests. In particular, memory demands (e.g., interference effects across study faces and/or across lineups) and task demands (e.g., not knowing for which study face a particular lineup had been created, as both study face presentation order and testing order were randomized) might have caused sequential participants to adopt a different decision strategy compared to when they are confronted with only a single lineup. In keeping with this idea, sequential participants chose foils and look-a-likes at a significantly higher rate in the random lineups used in our study ($M = .36$ and $M = .24$, respectively) compared to the meta-analysis reported by (Stebly et al., 2001) ($M = .19$ and $M = .09$, respectively). Though foil and look-a-like choice rates in our matched lineups were more comparable ($M = .13$ and $M = .30$, respectively) to the meta-analysis, the overall higher rates obtained suggest that sequential participants may have, say, chosen a face that looked similar to study face A, when in actuality the lineup they were viewing was for study face B. Though this could also occur in simultaneous lineups, it would be more problematic in sequential lineups because faces were being viewed one at a time. Thus, there are two possible issues: 1) the ideal sequential procedure was compromised by task demands, and 2) the task was easier in the simultaneous compared to sequential condition.

In view of these important considerations, the demands of the task are understood a bit further by considering a few more aspects of the results. Performance decrements across trials would indicate that fatigue, memory weakening, and/or interference were affecting the results. If present, performance decrements would suggest that the number of test trials influenced the decision strategy that is normally used when participants are confronted with a single lineup. However, analysis of identification outcomes by the order in which the lineups were conducted showed no performance decrements over the 12 trials for sequential or simultaneous lineups. In addition, the distribution of responses across lineup members in the sequential condition did not differ on the first and last trial, suggesting that sequential participants maintained a similar decision strategy across the 12 lineup tests.⁷

Furthermore, interference effects should have been greater in simultaneous lineups because participants saw more faces compared to the sequential condition. Simultaneous participants saw all 6 faces in each lineup, while sequential participants did not see the remaining photos after selecting a face. In addition, mock participants did not pick the suspect out at greater than chance levels from a simultaneous lineup, demonstrating that simultaneous witnesses did not have a guessing advantage by virtue of having all of the faces present. These results strongly suggest that the demands of the task were comparable between the two lineup procedures. Lastly, we obtained the standard pattern of findings found in comparisons of simultaneous and sequential lineups (Stebly et al., 2001), with hits and false alarms larger in simultaneous lineups, and lineup rejections larger in sequential lineups. Given this consistency, we do not think that the similarity effects that we obtained in sequential lineups arose because of the fact that 12 lineups were shown instead of one.⁸

Though it is premature to apply these results to actual cases given the short time frames and the limited similarity parameters used in the present study, the results demonstrate that identification decisions in simultaneous and sequential lineups are affected in much the same way by the similarity characteristics of the lineup. In both procedures, foil-study face similarity affected the decision criterion, and consequently, look-a-like identifications. Further research is needed to establish to what extent the look-a-like should match the culprit for the foil similarity effects on accuracy to take hold, and whether the effects are larger if the foils are matched on a greater number of features and/or characteristics (e.g., spatial relations among features). Finally, this work should consider the position of the suspect in the lineup, as position effects in sequential lineups may arise as a result of the similarity relations among members.

References

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

⁷ What is more, we also had participants rank order the similarity of the study face to the lineup members based on either their memory for the study face or while the study face was in view. Correspondence between the rankings and identification outcomes did not differ depending on whether the ranks were made based on memory or direct visual comparisons of the study face and the lineup members, indicating that the participants were able to retain the individual study faces in memory through all 12 lineups.

⁸ Still further, one might argue that the demands that we placed on our participants are more comparable to the demands faced by real world witnesses (who have to distinguish bystanders from culprit(s), who risk weakening of memory by seeing additional faces during the interval between the crime and lineup test, etc.) than the typical identification study, which presents participants with a single face and then tests after a 5-min retention interval, during which time, no other faces intervene (Flowe, Ebbesen, Burke, & Chivabundt, 2001). In capturing real world identification tests, the face validity (i.e., we showed more than one face, more than one lineup) of our study is low. However, in terms of the memory load that might actually be involved in the decision process, our study may be more comparable.

- Benjamin, A. S. (2005). Recognition memory and introspective remember/know judgments: Evidence for the influence of distractor plausibility on “remembering” and a caution about purportedly nonparametric measures. *Memory and Cognition*, 33(2), 261–269.
- Benjamin, A., & Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory and Language*, 51(2), 159–172.
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, 29(4), 395–424.
- Clark, S. E., & Davey, S. L. (2005). The target-to-foils shift in simultaneous and sequential lineups. *Law and Human Behavior*, 29(2), 151–172.
- Ebbesen, E. B., & Flowe, H. D. (2002). Simultaneous v. Sequential Lineups: What Do We Really Know? Retrieved January 9, 2006, from <http://www-psy.ucsd.edu/%7eeebbesen/SimSeq.htm>
- Flowe, H. D., Ebbesen, E. B., Burke, C., & Chivabundit, P. (2001). *At the scene of the crime: An examination of the external validity of published studies on line-up identification accuracy*. Presented at the American Psychological Society, Toronto, Canada.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110, 306–340.
- Laughery, K. R., Jensen, D. G., & Wogalter, M. S. (1988). Response bias with prototypic faces. In M. M. Gruneberg, R. Sykes, & P. Morris (Eds.), *Practical aspects of memory: Current research and issues* (pp. 157–162). Chichester: Wiley.
- Lindsay, R. C. L., Lea, J. A., Nosworthy, G. J., & Fulford, J. A. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology*, 76(6), 796–802.
- Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior*, 18, 527–541.
- Lindsay, R. C., Wallbridge, H., & Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science*, 19(4), 463–478.
- Luus, C. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior*, 15(1), 43–57.
- 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 and Cognition*, 33(5), 783–792. *People v. Cobb*, No. B170957, 2004, Cal. App. LEXIS 2088.
- Stebly, 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.
- Stebly, 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.
- Stebly, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, 21(3), 283–297.
- Tunnicliff, J. L., & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions? *Law and Human Behavior*, 24, 231–258.
- Wells, G. L. (2006). The removal-without-replacement effect for lineups. Manuscript in preparation.
- Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48, 553–571.
- Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentification. *Psychological Bulletin*, 88, 776–784.
- 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 and Human Behavior*, 22(6), 603–647.
- Wogalter, M. S., Burger, M. A., & Malpass, R. S. (1993). *Methods of constructing live and photographic lineups by police officers in the United States*. Paper presented at the International Conference on Face Processing, Wales, United Kingdom.
- Wogalter, M. S., Marwitz, D. B., & Leonard, D. C. (1992). Suggestiveness in photospread line-ups: Similarity induces distinctiveness. *Applied Cognitive Psychology*, 6(5), 443–453.
- Xu, M., & Bellezza, F. S. (2001). A comparison of the multimemory and detection theories of know and remember recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1197–1210.