An Examination of Simultaneous Lineup Identification Decision Processes Using Eye Tracking

Heather D. Flowe
School of Psychology
University of Leicester

Garrison W. Cottrell
Computer Science and Engineering
University of California, San Diego

Abstract

Decision processes in simultaneous lineups (an array of faces in which a "suspect" face is displayed along with foil faces) were examined using eye tracking to capture the length and number of times that individual faces were visually analyzed. The similarity of the lineup target face relative to the study face was manipulated, and face dwell times on the first visit and on return visits to the individual lineup faces were measured. On first visits, positively identified faces were examined for a longer duration compared to faces that were not identified. When no face was identified from the lineup, the suspect was visited for a longer duration compared to a foil face. On return visits, incorrectly identified faces were examined for a longer duration and visited more often compared to correctly identified faces. The results indicate that lineup decisions can be predicted by face dwell time and the number of visits made to faces.

An Examination of Simultaneous Lineup Identification Decision Processes Using Eye Movements

Cognitive psychologists have long been interested in finding measures that can index the quality of a memory report. In the eyewitness memory field, several markers have been employed to assess the accuracy of lineup identification, including decision confidence (e.g., Brewer & Wells, 2006), description accuracy (e.g., Pigott & Brigham, 1985), and response speed (e.g., Sporer, 1994). These markers might one day enable the estimation of eyewitness identification accuracy in actual criminal cases. The cognitive processes at work during an eyewitness identification test might also be established by using such measures.

In a simultaneous lineup, the person suspected by the police, who may or may not be the actual culprit, is presented alongside a number of foils (i.e., distractors that are known by the police to be innocent), and the eyewitness has to determine which one, if any, is the culprit. Currently, little is known about how participants allot their time in examining individual faces in a simultaneous lineup. Can the length of time that a face is examined at test index memory for the face, even when the eyewitness does not explicitly remember the face? Does the length and number of visits made to the faces in a lineup relate to differences in accuracy? The present study addressed these questions by testing participants with simultaneous lineups while their eye movements were recorded. Recently, researchers have begun to investigate decision processes in lineups using eye movements (Caspers, Betts, Chowdhry, Peterson, & MacLin, 2009; Flowe, 2010; Mansour, Lindsay, Brewer, & Munhall, 2009). We sought to replicate and extend these studies to further our understanding of cognitive processes in lineups.

Lineup Decision Processes

The most widely accepted theory of lineup decision processes is the relative versus absolute decision model put forward by Lindsay and Wells (1985). An absolute decision process involves comparing a lineup member to a memory representation of the culprit, whereas in a relative decision process, the lineup members are compared to each other, and the one that is the relatively best match to the memory representation of the culprit is chosen. These processes may be thought of as endpoints on a continuum of eyewitness behavior rather than as strict dichotomies (Charman & Wells, 2007; Wells, 1984). Identifications that are based mostly on an absolute rather than a relative decision strategy are associated with increased identification accuracy (Wells et al., 1998). Other theorists have similarly modeled eyewitness identification decisions as being comprised of both relative and absolute decision strategies (Clark, 2003; Pozzulo & Lindsay, 1999).

Other work has found that participants who report that they deliberated across lineup members are less accurate than those who report that they used an automatic recognition process (Dunning & Stern, 1994). Automatic processing characterizes identifications that are rapid and accompanied by little cognitive effort and feelings of automatic recognition (e.g., "His face just popped out at me."), whereas deliberative processing characterizes slower and more effortful identifications, whereby the participant engages in a process of elimination. Research consistently finds that the total amount of time that it takes to issue a lineup response tends to be relatively rapid when a positive identification is accurate rather than inaccurate (e.g., Brewer, Caon, Todd, & Weber, 2006; Brewer, Gordon, & Bond, 2000; Sporer, 1992, 1993, 1994; Weber, Brewer, Wells, Semmler, & Keast, 2004). These results suggest that participants who use an automatic rather than a deliberative decision strategy have a higher rate of identification accuracy. However, studies that have attempted to encourage the use of automatic processing by limiting the duration of exposure to a lineup have not found concomitant increases in decision accuracy (Brewer, Gordon, & Bond, 2000; Pozzulo, Crescini, & Lemieux, 2008). Brewer and colleagues (2000) concluded from their findings that the use of an automatic decision strategy does not cause identification accuracy to increase. Rather, participant self-reports of automatic decision processes in lineup identification may indicate that the memory representation for the positively identified face was strong. One of the primary purposes of the present study was to examine whether the degree to which a previously studied face is represented in memory can be indexed by the length of time that participants fixate on the face in a simultaneous lineup. A second major aim was to capture evidence of deliberative processing to further examine its relationship with identification accuracy.

The Use of Eye Tracking to Understand Face Processing

Mainstream face recognition research has found that face memory can be indexed with visual behavior. Eye movements are recorded during recognition to measure the length of time that participants fixate on the test faces. The assumption made by these studies is that there is a link between cognitive processes and the length of time that a person fixates in a visual scene (see Just & Carpenter, 1980). Eye tracking research has found that visual sampling varies as faces become increasingly familiar (Althoff et al., 1998; Ryan, Althoff, Whitlow, & Cohen, 2000); Ryan, Hannula, & Cohen, 2007). Studies such as these suggest that eye tracking can be used to study face memory in simultaneous lineup identification.

Eye tracking research with simultaneous lineups has found that face dwell time length (i.e., the total length of time that participants fixate on a face during a lineup trial) varies across lineup members in relation to identification response (Flowe, 2010; Mansour et al., 2009). Face dwell time is longer for a positively identified face than it is for any other face in a lineup. Dwell time length also tends to be longer for an incorrectly positively identified face compared to a correctly identified face. Mansour and colleagues (2009) found that fewer comparisons were made across faces when the target was correctly identified rather than a foil face. People were observed to engage in both non-exhaustive searches (indicative of an absolute strategy) and comparison behavior (indicative of a relative strategy), illustrating that the decision process in simultaneous lineups can be a combination of both types of processes. The majority of the time, however, comparison behavior was observed.

One challenge that arises in utilizing face dwell time to measure the memory for a lineup member is that face dwell time length measures several underlying cognitive operations. Dwell time length may reflect not only memory for a face, but also visual processing, error checking and deliberation. Additionally, when participants fixate on an individual face, they may also be determining whether the face is relatively more similar to the perpetrator compared to the other lineup members they have seen. With this in mind, the present study sought to capture memory for a face by measuring dwell time length for the first visit that was made to a face, as well as the length of time that a face was visited on subsequent visits. The rationale for measuring dwell time length on first and return visits is that relative comparison processes were assumed to play a more limited role in the visual analysis of a face the first time that the face is encountered in a lineup compared to subsequent visits made to the face. First visit and return visit dwell time lengths were then examined across lineup members to determine whether they were associated with face memory and identification responses.

Overview and Predictions of Present Study

Participants studied a series of faces and then attempted to identify each one from a simultaneous lineup while their eye movements were recorded. Dwell time length for each face was measured on first and return visits, and these measures were analyzed in relation to face type (foil or target) and identification response (pick target or foil, or reject lineup). Additionally, the similarity of the target to the study face was manipulated to determine whether dwell time length varied depending on the familiarity of the target face. The following hypotheses were tested:

Hypothesis 1: First visit dwell time length will be longer for target faces compared to foil faces, regardless of identification response.

If dwell time length on the first visit measures memory for a face, then differences are expected in how long people look at faces depending on whether they were previously studied. Note that this hypothesis is testing whether first visit dwell time length may be used as an implicit memory marker, as longer first visit dwell times are predicted regardless of identification response. Longer dwell times for the target compared to the other faces would be expected if first visit dwell time indexes face familiarity even in the absence of conscious recollection. On the other hand, if longer first visits to target compared to foil faces are obtained only when the target is identified, this would indicate that eye movements are directed toward a stimulus only when the stimulus is consciously recollected.

Longer first visits were predicted for targets compared to foils because previous eye tracking research in simultaneous lineups has found that target faces are attended to for a longer period compared to foils (Flowe, 2010; Mansour et al., 2009). Sequential sampling models of recognition also lead to this prediction, as research has shown that participants are able to rapidly determine when a stimulus is not familiar (Link & Heath, 1975; Ratcliff, 1978; Van Zandt, 2000). Applied to lineup identification, participants know they have to rule out most of the test faces, as only one face, if any, can be identified. Therefore, participants were expected to more rapidly determine that a face was unfamiliar rather than familiar, which would be indexed by shorter first visits for foil compared to target faces.

Hypothesis 2: The more similar a target face is to the study face, the longer the dwell time length will be on the first visit, regardless of identification response.

Target face similarity with respect to the study face was expected to have an association with first visit length. The more similar a face is to the study face, the more familiar it should seem. Consequently, people will look at the target face increasingly longer as the target becomes more similar to the study face.

Hypothesis 3: Return visit dwell time length for a positively identified face will be longer if the positively identified face is a foil rather than a target.

As previous research suggests that simultaneous lineup identifications are a product of both automatic and deliberative processing (Mansour et al., 2009), we expected that participants would return to the most familiar face in the array for a longer length of time to determine whether it should be identified and deliberate longer when their memory was in error.

Hypothesis 4: Inaccurate identifications will be accompanied by a greater number of face visits than accurate identifications.

If return visits capture deliberation processes that arise when memory for the study face is relatively weak, then the number of visits made to a positively identified face should be greater when a foil rather than the target is identified from the lineup.

Method

Participants and Study Design

Thirty-nine undergraduates (mean age=21.10, 22 female) participated in the study for course credit. All had normal or corrected-to-normal visual acuity. The independent variables were measured within subjects, and included target condition (whether the target or the look-alike was present in the lineup), and face type (whether the lineup member under examination by the subject was a non-identified, or rejected, foil, a positively identified foil, or the suspect (i.e., target or look-alike).

Stimuli and Apparatus

Face stimuli were created using FACES 4.0, a composite drawing software program used by law enforcement. A database of 1000 faces was created by having the program generate faces at random. None of the faces had distinguishing features, facial hair, or head hair. We ensured for every participant that none of the faces in the study set and in the lineup sets shared any features. The configuration of the facial features across the lineup members, however, was held as constant as possible, thereby leaving the internal features of the face to vary across members. These stimuli were utilized in order to exercise tight experimental control over the similarity structure of the faces. By eliminating extraneous variation in the faces, which is added by hairstyle and configural differences, the statistical power to detect differences in response time across target and foil faces should be increased. Of course, exercising tighter control over the properties of the faces reduces the range of circumstances to which these findings may be applied. However, for the sake of theory building, experimental control is essential. For instance, much has been learned about the workings of memory by utilizing tightly controlled stimuli, such as word lists. These studies serve as a baseline for subsequent studies that are then able to examine how the complexity of the to-be-learned information impacts the pattern of results. Additionally, a number of important studies that have revealed the inner-workings of face processing have been conducted using composite drawings of faces (e.g., Leder & Bruce, 2000; Tanaka & Farah, 1993; Tanaka & Sengco, 1997). Arguably, the faces in many of these studies are more artificial looking than that those produced by FACES. We have replicated effects with FACES that have been obtained with more naturalistic faces, including inversion effects (Flowe, 2010) and the sequential lineup advantage (Flowe & Ebbesen, 2007).

Six person lineups were formed by randomly selecting without replacement 6 faces from the database. From these 6 faces, the to-be-identified study face was randomly chosen. Using this procedure, 12 study faces and 12 accompanying lineups were established. 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-alike") was created for every lineup. This new version of the study face was obtained by substituting one feature from the original study face (eyes, nose or mouth) for another; the configuration (distance between the eyes and eye height, location of the mouth and nose) was held constant. The test faces were presented together in a 3 X 2 array with a number placed beneath each picture for purposes of identification. For each participant, the identical target was present in the array on half of the trials, and the look-alike was present on the other half. The study faces were 15 cm in height and the test faces were 8 cm in height (the width of the faces was constrained by the natural proportions of the face). Please see Figure 1 for an example of a lineup.

Eye movements were recorded by an Eyelink II (SR Research) video-based eye tracker with a temporal resolution of 250 Hz and a spatial resolution of 0.2 degrees; the eye tracker default settings for cognitive research were used. Data from the eye with less error during calibration were used for analysis. Participants sat approximately at a distance of 50 cm away from the monitor. Hence, the study faces subtended about 17 degrees of visual angle, and each of the test faces 9.7 degrees of visual angle.

Procedure

Participants were told that they would be given a list of faces to study and to pay careful attention because they would have to recognize them on a later test. Following a calibration phase, the 12 study faces were presented in succession for 30 s each, with an interstimulus interval of 3 s. The delay between the study period and the test period was 5 min; during the delay, the eye tracker was recalibrated.

Participants were instructed that only one face, if any, could be identified from a given lineup. They were also told that the lineup should be rejected if none of the faces was a perfect match to one of the study faces. Each identification trial began with a drift correction wherein a central fixation point was displayed on the screen followed by the identification test. No response deadline was imposed. Participants verbally indicated their response to the experimenter, who then entered the answer using a keyboard. The recording of eye movements was synchronized with the keyboard press; hence, as soon as participants indicated their response, eye tracking for the trial was terminated.

Dependent Measures and Data Analysis

The dependent variables included: identification outcome, total face dwell time, first visit dwell time, return visit dwell time, and number of face visits. Identification outcome (lineup rejection rates and the rates of identifying a foil, the look-alike, and the identical target) were computed across trials for each participant. Total dwell time was calculated for each participant by summing across the fixation times that were obtained for each face in a given lineup test trial.

Total dwell time was conceived of as a corollary to response time, which has been used in previous lineup identification research. Face dwell times were then averaged across trials, conditioning the data on face type (target, look-alike, or foil), target condition (identical target or look-alike present), and identification outcome (identify target, look-alike, or foil, or reject lineup). First visit dwell time was determined by summing across the fixation times that occurred on the first visit to the face (i.e., from the first time the face was encountered at test until the time when the participant fixated his or her eyes on another face in the display). Return visit dwell time was determined by summing across all of the fixation times made to the face from after the first visit was made to the face until the end of the lineup trial (Also see Figure 1.). The number of visits made to each face was also obtained. A visit to a face concluded when another face was fixated.

Lastly, the similarity of each look-alike faces compared to their respective study face was established by having a group of raters (N=30) rate their similarity using a 101 point scale, anchored at 0, “not at all similar,” and 100, “completely identical.” On average, the look-alike faces were rated 71.25 (SD=10.93) in their similarity to the study face. The similarity ratings were correlated with first visit dwell times to test the hypothesis that first visit dwell times would increase along with target face familiarity.

The dependent measures were analyzed with ANOVAs and t-tests; partial eta-squared (η_p^2) is provided as a measure of effect size. Dwell time data were positively skewed, and therefore, were square root transformed before submitting them to inferential statistical analysis. The pattern of results was the same for the transformed and untransformed scores; therefore, descriptive statistics are based on the untransformed data for interpretation ease.

Results

Identification Responses

Participants were more likely to identify the target compared to the look-alike, and the difference was significant, $F(1, 38)=12.53$, $p<.05$, $\eta_p^2=.25$ (see Table 1). Replacing the target with the look-alike resulted in a higher rate of rejection responses. The lineup rejection rate was higher when the look-alike rather than the target was in the lineup, $F(1, 38)=19.92$, $p<.05$, $\eta_p^2=.34$. The foil identification rate, however, did not significantly vary across target condition.

Total Lineup Response Times and Accuracy

Identification accuracy was analyzed in relation to the total amount of time spent dwelling on all of the faces during the trial. The goal was to replicate (using total trial dwell time for each face) the typical finding that there is a negative relationship between lineup response time latency and positive identification accuracy. In keeping with previous research, the total amount of time spent dwelling on faces in target present lineups was shorter when the positive identification was correct (M=11.8 s) rather than incorrect (M=15.3), $F(1, 25)=5.47$, $p<.05$, $\eta_p^2=.18$. When the look-alike was present, the average total dwell time obtained for the trials in which a foil was positively identified (M=16.2 s) did not reliably differ from the trials in which the look-alike was identified (M=12.2 s). With regard to lineup rejections, the average total dwell time for correct rejections when the look-alike was present (M=14.2 s) did not differ from incorrect rejections of the lineup when the target was present (M=13.7 s).

First Visit Face Dwell Times

Table 2 displays first visit dwell time data by face type, target condition, and identification outcome. Additionally, Figure 1 (top panel) illustrates first visit dwell time length on a typical trial. Average first visit dwell time for the identical target compared to the look-alike was compared within every identification outcome. First visit dwell time was marginally higher for the identical target when the target was positively identified ($t(34)=1.68$, $p=.05$). Identical target and look-alike dwell times did not differ when the lineup was rejected ($t(27)=.98$, $p=.33$), or when a foil was positively identified ($t(21)=.84$, $p=.41$). Therefore, the look-alike and target data were collapsed in the analyses that follow. Positive identifications of the identical target or the look-alike will be referred to as *target identifications*, and visits to the identical target or the look-alike will be referred to as *target visits*.

Hypothesis 1 posited that first visit duration should be longer for the face in the lineup that was previously studied if first visit dwell time length measures memory for a face, even in the absence of conscious recollection. This hypothesis may be tested by examining whether target first visit dwell times are greater than the average length of time that a foil face was examined. Collapsed across ID responses, the target was indeed visited for a longer duration compared to a foil on the first visit, $F(1, 32)=11.06$, $p<.05$, $\eta_p^2=.26$. The data were next analyzed within each outcome to more directly test Hypothesis 1. The results were consistent with the hypothesis in target ID trials and in lineup rejection trials. On the first visit, the average dwell time for a non-identified foil compared to the target was shorter when the target was identified, $t(38)=5.78$, $p<.01$, and when the lineup was rejected, $t(36)=2.59$, $p<.05$. The results for foil identification trials were not consistent with the hypothesis, however. When a foil was positively identified, first visit dwell times for the non-identified target and a non-identified foil face did not differ, $t(32)=.57$, $p>.05$. Therefore, Hypothesis 1 was not supported.

The results thus far suggest that the face in the array that was perceived by the participant to be the most familiar was examined the longest on a first visit. If true, then positively identified faces should be visited longer than the other lineup members on a first visit, even if the participant was going to mistakenly identify a foil. The results were in keeping with this proposition. Positively identified faces were examined longer on a first visit compared to non-identified faces, $t(36)=5.77$, $p < .01$. Moreover, first visit dwell times for the positively identified face did not differ depending on whether the target rather than a foil was positively identified, $t(36)=-0.28$, $p>.05$. In positive identification trials, first visit dwell time length predicted which lineup member was chosen nearly half of the time. That is, in 45% of target identification trials, first visit length was the longest for the target when compared to all of the other lineup members. Similarly, in 41% of the trials in which a foil was positively identified, the positively identified foil was visited the longest when compared to all of the other lineup members.

The length of time that the target was first visited also varied with identification outcome. First visit dwell times for the target were longer if the target was positively identified compared to when the lineup was rejected, $t(36)=2.66$, $p<.05$, and when compared to foil identification trials, $t(32)=5.50$, $p<.05$. Interestingly, target first visit dwell time length was also longer when the lineup was rejected compared to when a foil was positively identified, $t(31)=2.49$, $p<.05$. This result suggests that the target was perceived as familiar on some trials, even when he was not positively identified. Further analysis found that the target was visited longer than any of the other lineup members on 37% of lineup rejection trials.

Hypothesis 2 stated that target similarity would be positively related to first visit dwell time. Though dwell time length was longer for the identical target compared to the look-alike, the difference was only marginally statistically significant. However, there was a statistically significant positive association between first visit dwell time length and look-alike similarity to the study face, r=.36, $p < .001$. First visit dwell time increased as the similarity of the look-alike compared to the study face increased, a finding that is in keeping with Hypothesis 2.

Return Visit Dwell Times

Table 2 displays return visit dwell time data conditioned on face type, target condition, and identification outcome. Additionally, Figure 1 (bottom panel) illustrates return visit dwell time length on a typical trial. Target condition did not have a significant effect on return visit dwell times, nor did it interact with any of the other variables; therefore, the look-alike and identical target data were collapsed in the analysis.

Hypothesis 3 stated that return visit dwell time length would be longer for a positively identified foil face compared to a target face. Results did indicate that the identified target was visited for a shorter duration compared to an identified foil, $t(32)$=1.97, p<.05. Additionally, non-identified foils were examined longer when a foil was positively identified compared to the trials in which the target was positively identified, $t(31)$=3.02, p<.05. These results are consistent with the hypothesis that participants deliberate more when they incorrectly identify a foil face.

Results further indicated that return visit duration was associated with identification outcome, $F(1, 22)$=111.67, p<.05, η_p^2=.83. Return visit target dwell times were longer in target identification trials compared to lineup rejection trials, $t(35)$=4.30, p<.05, and compared to foil identification trials, $t(24)$=7.01, p<.05. Additionally, return visit duration for the target was longer in rejection trials compared to foil identification trials, $t(22)$=2.95, p<.05.

In rejection trials, participants appeared to shift their attention from the target to a foil face in making return visits. Specifically, return visit duration for a foil was longer in rejection trials compared to target identification trials, $t(36)=7.02$, $p<.05$, and when compared to foil identification trials, $t(30)=2.08$, $p<.05$.

Number of Face Visits

The number of trials in which every face in the lineup was inspected before the first return visit to a face was made was determined for every participant. Participants were remarkably systematic in how they proceeded: On average across participants, every face was inspected before the first return was made on 82% of the trials.

Table 3 displays average number of face visits by identification outcome, face type, and target condition. Number of visits did not vary depending on whether the face was the target or the look-alike; therefore, we collapsed across target condition.

Hypothesis 4 posited that a greater number of visits would be made to a positively identified foil face compared to when a target was identified. The results indicated that a positively identified foil face was visited a greater number of times compared to a positively identified target, $t(32)=2.91$, $p<.05$. This appeared to be a fairly common response pattern across participants: We determined for every participant that had made both foil and target identifications the average number of times they visited a positively identified foil, and the average number of times that they visited a positively identified target. For 67% of participants, the average number of visits that were made to a positively identified foil face was greater than the average number of visits that were made to a positively identified target face. This suggests that deliberation typically occurs when the memory strength signal for a face is low, which would be expected for a foil face that is incorrectly identified.

Additionally, a greater number of visits were made to non-identified foil faces when a foil was positively identified compared to when the target was identified, $t(32)=2.39$, $p<.05$. Visits to non-identified foil faces were also more numerous when the lineup was rejected compared to when a foil was positively identified, $t(31)=2.03$, $p<.05$.

Discussion

Using measures derived from eye tracking, the finding that response latency and accuracy are negatively related on positive identification trials was replicated (e.g., Flowe, in 2010; Mansour et al., 2009; Sporer, 1992). We further found that first visit dwell time postdicted which of the faces was ultimately positively identified; longer first visit dwell times were obtained for positively identified faces. These results held regardless of whether the positively identified face was the target. Therefore, contrary to Hypothesis 1, it does not appear that dwell time can index implicit memory for a simultaneous lineup face. Rather, first visit dwell time is longer for faces that are consciously recollected as being familiar, even when recollection is in error. Dwell times were also higher the more similar the target was to the study face, a result that is in keeping with Hypothesis 2. The eye movement data also captured visual behavior on incorrect positive identification trials that might be indicative of deliberative processing, thereby supporting Hypothesis 3 and Hypothesis 4. The length and number of return visits made to the face that was ultimately positively identified were higher if the face was a foil rather than the target.

The familiarity of the lineup member was positively associated with face dwell time length on first and return visits. A foil face was examined for a shorter length of time than the target face. Additionally, dwell times for the look-alike compared to the identical target were consistently lower across all conditions, though the difference was not statistically significant. However, a relationship was found between the degree of look-alike similarity and visual behavior. Dwell times increased as the similarity of the look-alike with respect to the study face increased. These results indicate that participants examine a test face for increasing longer the more familiar the face seems to be. Interestingly, first visit dwell times were longer on average for the target face in lineup rejection compared to foil identification trials, further suggesting that first visit dwell time can index which of the faces in the array is the one that is most familiar to the participant, even if the face is not ultimately positively identified. Sequential sampling models of recognition posit that people appear to assess not only the familiarity of a stimulus but also its novelty (Link & Heath, 1975; Ratcliff, 1978; Van Zandt, 2000), a proposition that now has support from more recent studies in neuroscience (e.g., Viskontas, Knowlton, Steinmetz, & Fried, 2006). The foil faces, which should be the least familiar members of the lineup by virtue of them not having been studied, are rapidly rejected for identification compared to a target face. Another inference that might be made from the results is that a face will be identified if it surpasses the threshold for a positive identification. If the face falls below threshold, however, the lineup will be rejected or comparisons across lineup members are made. Additional research is needed to learn more about the types of decision strategies that arise when a face falls below threshold. For example, participants may set two thresholds, one for positive identifications and one for rejecting a face, and if a test face falls between the two thresholds, deliberation may occur (also see Clark, 2003).

The results further suggested that the tendency for people to engage in deliberative processing is not solely attributable to individual differences in lineup identification strategies. Rather, there appears to be a general tendency to deliberate when the match between a test face and a memory representation is relatively weak, as two-thirds of the sample tended to visit a positively identified face more often when they were positively identifying a foil rather than the target. This raises the question of what it is that participants are doing when they make multiple visits to a face under conditions in which memory strength for the face is relatively low. Dunning and Stern (1994) proposed that people may be devising decision rules when they deliberate, which in turn increases response time. The methodology of the present study does not allow for addressing this issue further. Additional research is needed to understand the cognitive operations in which participants engage when deliberating across lineup members.

The results of the present study are in keeping with the proposition that simultaneous lineup identifications are arrived at through a combination of absolute and relative decision processes (see Clark, 2003; Pozzulo & Lindsay, 1999), a conclusion that was also reached by Mansour and colleagues (2009) in their simultaneous lineup eye tracking study. The first visit dwell time results obtained in the present study suggest that one face appeared more familiar than the other faces on the first visit. Participants then returned to this face for a longer length of time than less familiar faces. At this point during the identification trial, correct and incorrect positive identifications could be differentiated. Participants returned to the correctly identified face for a shorter length of time and made fewer visits to the correctly identified face compared to trials where a face was incorrectly identified. In other words, accurate recognition appeared to be more automatic and less deliberative, a result that is in keeping with Brewer et al.'s (2000) proposal that automatic lineup decisions arise when there is a strong match between a test face and a visual memory.

The generalizability of the results is potentially limited by a number of factors. First, several identification tests were given to each participant. Actual eyewitnesses are typically confronted with only one lineup test (Flowe, 2007). In the present study, several trials were given to a research participant so that the basic cognitive processes in simultaneous lineups could be explored. Having said that, studies that have given participants multiple lineup trials have replicated the findings of studies in which a single lineup test is given to participants (Flowe & Ebbesen, 2007; Meissner et al., 2005), which suggests that decision processes may be studied under conditions in which participants are given multiple lineup tests. Another factor that limits the generalizability of our work is that we used tightly controlled face stimuli. Therefore, the methodology that was utilized in the present study shares more in common with other basic research on cognitive mechanisms. Finally, we agree with Mansour and colleagues (2009) that the effects observed in eye movement data are small and variable. Features of the to-be-remembered event and the lineup faces can affect response time, thereby making it inappropriate to say that there is an ideal amount of time that distinguishes an accurate identification over an inaccurate one (Weber et al., 2004). For all of these reasons, it would be premature to try to apply these findings to actual legal cases without performing many additional studies that employ a range of conditions that are more similar to the ecology of actual eyewitnesses.

In summary, first and return visits that are made to faces appear to be a promising marker for assessing lineup identification accuracy. We replicated the finding that there is a negative association between positive identification accuracy and response latency. It was further demonstrated that first visit dwell times can postdict identification outcomes; the longer the dwell time to a face on a first visit, the more likely it will be identified. Return visit dwell time and the number of visits made to a face differentiated correct from incorrect positive identification outcomes. Incorrect positive identifications are characterized by longer and more frequent return visits. Additional research is needed to further examine first and return visits to faces as potential memory markers for elucidating decision processes in lineups.

References

Althoff, R. R, & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. Journal of Experimental Psychology: *Learning, Memory, & Cognition, 25*, 997-1010.

Brewer, N., Caon, A., Todd, C., & Weber, N. (2006). Eyewitness identification accuracy and response latency. *Law and Human Behavior, 30*(1), 31-50.

Brewer, N., Gordon, M., & Bond, N. (2000). Effect of photoarray exposure duration on eyewitness identification accuracy and processing strategy. *Psychology, Crime & Law, 6*(1-2), 21-32.

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*(1), 11-30.

Charman, S., & Wells, G. L. (2007). Applied lineup theory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of Eyewitness Psychology: Memory for People*, Vol. 2. (pp. 219-256). NJ: Erlbaum.

Caspers, H., Betts, R., Chowdhry, O., Peterson, D., & MacLin, O. (2009). Uncovering attentional differences in simultaneous and sequential lineups using eye tracking. Paper presented at the American Psychology-Law Society, San Antonio, TX.

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

Dunning, D., & Stern, L. B. (1994). Distinguishing accurate from inaccurate eyewitness identifications via inquiries about decision processes. *Journal of Personality and Social Psychology, 67*(5), 818-835.

Flowe, H. D. (2007). Testing the reflection assumption: An examination of the external validity of lineup identification research. Off the Witness Stand: Using Psychology in the Practice of Justice. New York, NY.

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.

Flowe, H. D. (2010). An Exploration of Visual Processing Strategies in Eyewitness Identification Tests. *Applied Cognitive Psychology*, in press.

Flowe, H. D. (2010). Composite and inversion effects examined in simultaneous and sequential lineups: Is there a difference between procedures in configural/holistic processing? Manuscript in preparation.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

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.

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.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*, 77-105.

Mansour, J., Lindsay, R. C. L., Brewer, N., Munhall, K. G. (2009). Characterizing visual behavior in a lineup task. *Applied Cognitive Psychology*, *23*(7), 1012-1026.

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.

Pigott, M., & Brigham, J. C. (1985). Relationship between accuracy of prior description and facial recognition. *Journal of Applied Psychology, 70*(3), 547-555.

Pozzulo, J., Crescini, C., & Lemieux, J. (2008). Are accurate witnesses more likely to make absolute judgments? *International Journal of Law and Psychiatry, 31*(6), 495-501

Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child witnesses. *Journal of Applied Psychology, 84*, 167-176.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review, 85*(2), 59-108.

Ryan, J., Hannula, D., & Cohen, N. (2007). The obligatory effects of memory on eye movements. *Memory, 15*(5), 508-525.

Ryan, J., Althoff, R., Whitlow, S., & Cohen, N. (2000). Amnesia is a deficit in relational memory. *Psychological Science, 11*(6), 454-461.

Sporer, S. L. (1992). Post-dicting eyewitness accuracy: Confidence, decision-times and person descriptions of choosers and non-choosers. *European Journal of Social Psychology, 22*(2), 157-180.

Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78*(1), 22-33.

Sporer, S. L. (1994). *Decision times and eyewitness identification accuracy in simultaneous and sequential lineups*. In D.F. Ross, J.D. Read & M.P.Toglia (Eds.), Adult eyewitness testimony: Current trends and developments (pp. 300-327). New York: Cambridge University Press.

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.

Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 582-600.

Viskontas, I. V., Knowlton, B. J., Steinmetz, P. N., & Fried, I. (2006). Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *Journal of Cognitive Neuroscience, 18*(10), 1654-1662.

Weber, N., Brewer, N., Wells, G. L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: The unruly 10-12-second rule. *Journal of Experimental Psychology: Applied, 10*(3), 139-147.

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.

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14*, 89-103.

Table 1

Average Rate for Identification Outcomes (*SE*) by Identification Procedure and Target Condition.

	Suspect	Foil	Reject
Target Present	0.48 (.04)	0.24 (.04)	0.28 (.04)
Target Absent	0.32 (.04)	0.19 (.03)	0.49 (.04)

Table 2

Average First and Return Visit Dwell Times in ms by Face Type (Target or Foil), Identification (ID) Response and Target Condition (Identical or Look-alike in the Lineup).

ID Response	**First Visit Dwell Time (*SE*)**			**Return Visit Dwell Time (*SE*)**		
Face Type	Identical	Look-alike	Overall	Identical	Look-alike	Overall
Target/Look-alike ID						
Target	1109 *(112)*	1004 (109)	*1067* *(99)*	3146 *(478)*	3214 *(566)*	3174 *(457)*
Foil	640 *(47)*	590 (48)	*620* *(41)*	638 *(97)*	508 *(93)*	586 *(88)*
Foil ID						
Target	576 *(87)*	585 *(61)*	*580* *(61)*	951 *(197)*	1264 *(358)*	1086 *(211)*
Identified Foil	993 *(85)*	1005 (88)	*998* *(64)*	3618 *(770)*	3809 *(1094)*	3701 *(829)*
Non-Identified Foil	625 *(46)*	626 *(41)*	*625* *(33)*	844 *(171)*	1100 *(428)*	956 *(247)*
Lineup Rejected						
Target	1023 *(102)*	798 *(69)*	*881* *(67)*	2125 *(350)*	2123 *(304)*	2124 *(282)*
Foil	669 *(47)*	633 *(40)*	*647* *(36)*	1201 *(221)*	1366 *(179)*	1306 *(178)*

Table 3

Average Number of Face Visits (*SE*) by ID (Identification) Outcome, Target Condition, and Face Type.

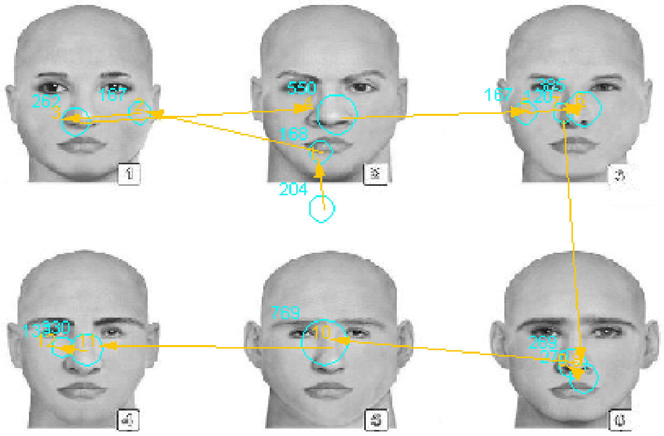
ID Response / Face Type	Number of Visits (*SE*) Present	Absent	*Overall*
Target/Look-alike ID			
Target/Look-alike	2.80 (*.23*)	3.00 (*.22*)	*2.78 (.17)*
Foil	2.07 (*.15*)	2.05 (*.22*)	*2.02 (.15)*
Foil ID			
Target/Look-alike	2.41 (*.16*)	2.36 (*.32*)	*2.33 (.19)*
Identified Foil	3.48 (*.23*)	3.26 (*.32*)	*3.46 (.21)*
Non-Identified Foil	2.38 (*.24*)	2.08 (*.23*)	*2.38 (.21)*
Lineup Rejected			
Target/Look-alike	3.05 (*.31*)	2.87 (*.24*)	*2.97 (.21)*
Foil	2.70 (*.23*)	2.56 (*.19*)	*2.70 (.18)*

Figure Captions

Figure 1. Illustration of a lineup and a typical test trial. Fixations (circles) and saccade pattern (lines) are displayed for the first visit (top panel) and return visits (bottom panel) made to each face. Fixation length is captured by the relative size of the circle. First visit dwell time length for a given face was determined by summing across all of the fixations that were made to the face on a first visit. Return visit dwell time length was determined for a given face by summing across all of the fixations that were made to the face on a return visit.

Figure 1

First Visit



Return Visit