Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis

CHRISTIAN A. MEISSNER
Florida International University, Miami, Florida

COLIN G. TREDOUX
University of Cape Town, Cape Town, South Africa

JANAT F. PARKER
Florida International University, Miami, Florida

and

OTTO H. MacLIN
University of Northern Iowa, Cedar Falls, Iowa

Many eyewitness researchers have argued for the application of a sequential alternative to the traditional simultaneous lineup, given its role in decreasing false identifications of innocent suspects (*sequential superiority effect*). However, Ebbesen and Flowe (2002) have recently noted that sequential lineups may merely bring about a shift in response criterion, having no effect on discrimination accuracy. We explored this claim, using a method that allows signal detection theory measures to be collected from eyewitnesses. In three experiments, lineup type was factorially combined with conditions expected to influence response criterion and/or discrimination accuracy. Results were consistent with signal detection theory predictions, including that of a conservative criterion shift with the sequential presentation of lineups. In a fourth experiment, we explored the phenomenological basis for the criterion shift, using the remember–know–guess procedure. In accord with previous research, the criterion shift in sequential lineups was associated with a reduction in familiarity-based responding. It is proposed that the relative similarity between lineup members may create a context in which fluency-based processing is facilitated to a greater extent when lineup members are presented simultaneously.

Wrongful conviction of the innocent is likely the most egregious error that occurs in the criminal justice system. Yet, through the recent advent of DNA testing, exonerations of the innocent are increasing at a rate that few in the criminal justice system would have predicted (Scheck & Neufeld, 2001). Studies in which cases involving wrongful conviction have been examined have consistently indicated that mistaken eyewitness identification evidence is the leading contributing factor (Huff, Rattner, & Sagarin, 1996; Scheck, Neufeld, & Dwyer, 2000). However, findings that mistaken eyewitness identifications frequently lead to wrongful conviction come as no surprise to psychologists who study eyewitness memory (Wells & Olson, 2003). Over the past 20 years, eyewitness researchers have pursued various procedures that

might reduce such mistaken identifications in everyday practice.

A much heralded procedural innovation in eyewitness identification was suggested by Lindsay and Wells (1985) and subsequently investigated by Lindsay and colleagues in more than a dozen studies (see Lindsay, 1999). These researchers proposed that the lineup procedure typically used by law enforcement, in which all photographs are presented to the witness at the same time in a *simultaneous* format, may lead to increased false positive choices as a result of a particular face's being selected on the basis of that individual's familiarity relative to that of others in the lineup. They recommended that an alternative procedure might reduce this reliance upon relative judgments—namely, presenting photographs to the witness one at a time in a *sequential* format. Lindsay and Wells proposed that such a procedure would encourage the use of what they termed "absolute judgments," based largely on evaluating each face individually.

Since the original Lindsay and Wells (1985) study, a small body of research has accumulated that has demonstrated certain advantages to utilizing such a sequential lineup procedure. In a recent meta-analysis of these stud-

ies, Steblay, Dysart, Fulero, and Lindsay (2001) concluded that the sequential lineup does indeed significantly reduce the rate of false identifications when the target is absent from the array, in comparison with performance in the simultaneous lineup (i.e., the *sequential superiority effect*). However, this reduction occurs at the cost of a lower rate of correct identifications in the sequential lineup when the target is present in the array. Thus, the sequential procedure appears to reduce choosing, regardless of the presence or absence of the target face.

One interpretation of such a pattern has been provided by Ebbesen and Flowe (2002). They argued that the sequential lineup may raise participants' *response criterion*; that is, individuals may increase their threshold for deciding whether to respond that a stimulus has been seen before when faces are viewed in isolation from one another. The notion of a change in response criterion (or *criterion shift*) is based on Green and Swets's (1966) classic signal detection theory (SDT). Generally speaking, SDT provides a model for understanding how individuals recognize (or distinguish) certain objects or events that they have experienced previously from those that are novel. As such, SDT examines the types of decisions that individuals can make regarding old and new experiences, including their ability to correctly recognize an old experience (a hit response) and to falsely recognize a new experience as old (a false alarm response). By isolating these two decisions, SDT separates an individual's performance into two independent parameters—namely, *discrimination accuracy* (the ability of an individual to correctly detect a signal vs. correctly reject its absence) and *response criterion* (the degree of evidence necessary for the individual to respond that a signal has been presented). Generally speaking, factors that influence the quality of the memory representation (such as the time permitted to study/encode a stimulus) should influence discrimination accuracy. On the other hand, an individual's response criterion is influenced by various social or instructional factors that may create a proclivity or bias to respond in one manner or another (such as reward for the number of correct items recognized).

Although a criterion shift explanation provides a useful description of potential decision processes that may occur when simultaneous versus sequential arrays are responded to, it is also important to understand the phenomenology and memory processes that may be affected by such procedural manipulations. Dual-process theories of recognition memory have been widely advocated for their ability to account for the cognitive processes underlying performance (Gardiner & Richardson-Klavehn, 2000; Kelley & Jacoby, 2000). Generally speaking, such theories have distinguished between the retrieval of conscious-level conceptual information that is elaboratively encoded (*recollection*) and the use of fluent, perceptually based information that is believed to be encoded in an automatic, nonconscious manner (*familiarity*). In his review of the dual-process memory literature, Yonelinas (2002) demonstrated that recollection is generally influenced by such manipulations as generative or semantic

encoding, division of attention, speed of responding, and novel learning and is sensitive to the effects of aging and amnesia. In contrast, familiarity is often sensitive to fluency manipulations, forgetting over short retention intervals, and, most important for the present investigation, changes in response criterion.

Given the association between familiarity-based responding and fluctuations in response criterion, a dual-process theory of memory may assist us in understanding the *relative judgment* process believed to occur with procedural lineup manipulations. More specifically, perceived familiarity may be relied upon more in simultaneous arrays, given that presenting faces at the same time can provide a contextual basis for the application of familiarity or fluency. In contrast, reliance upon familiarity may be reduced when faces are presented independently from one another, since the contextual basis that facilitates such judgments is taken away.

In the present experiments, we sought to explore the criterion shift explanation offered by Ebbesen and Flowe (2002) for the sequential lineup and to investigate whether such procedural lineup manipulations might be understood within the context of a dual-process theory of memory. Unfortunately, the typical eyewitness identification paradigm prevents a direct application of SDT, primarily due to the fact that participant witnesses are exposed to a single event and target face and are subsequently asked to provide only a single identification decision involving a target-present or a target-absent lineup. In contrast, SDT parameters of discrimination accuracy and response criterion rely upon multiple identification decisions under conditions in which the target is sometimes present and sometimes absent (thereby producing estimates of hits and false alarms, respectively).

To remedy this problem, we introduced a novel *lineup recognition paradigm*. Participants were presented with multiple target faces and later attempted to identify these faces from a series of target-present and target-absent lineups. SDT parameters were then computed for each participant's performance, to examine the effects of various lineup manipulations. In addition to examining lineup presentation method, we investigated the effects of encoding strength, conservative lineup instructions, and nominal lineup size. These additional manipulations were explored in the present experiments because of their practical influence and to further estimate their influence on SDT measures on the basis of the patterns of influence shown in previous eyewitness studies. In a final experiment, we also obtained phenomenological judgments from participants, using the remember–know–guess procedure (see Gardiner & Richardson-Klavehn, 2000) to assess the contributions of recollection versus familiarity to eyewitness decisions in simultaneous and sequential lineup arrays.

GENERAL METHOD

Materials

A computer program developed by MacLin and colleagues (MacLin, Meissner, & Zimmerman, 2005; MacLin, Zimmerman, &

Malpass, 2005), PC_Eyewitness, was used to control the presentation of stimuli and the collection of data across all the experiments. The software was designed specifically for the presentation of eyewitness lineup arrays and has been shown by MacLin and colleagues to produce results consistent with that of the standard paper-and-pencil methodology used by the majority of eyewitness researchers. PC_Eyewitness is available to potential users free of charge and may be downloaded from http://www.uni.edu/~maclino/software.

The facial stimuli used in the present experiments were taken from a database maintained by the first author. Photographs presented at study involved head-and-shoulder views of male targets wearing their everyday clothes, standing in front of a gray background, in a full frontal, smiling pose. Photographs presented at test involved head-and-shoulder views of the targets wearing a burgundy-colored sweatshirt, standing in front of a gray background, in a full frontal, nonsmiling pose. Eight males were selected from the database for use as target faces, with selection based largely on ensuring variability in the facial features of target faces (hair color, eye color, shape of face, distinctive features, etc.). Twenty participants were then recruited to provide descriptions of each face, which were combined to create a modal description for each target face. This modal description was used as a basis for selecting fillers to compose 6-person (Experiments 1, 2, and 4) and 12-person (Experiment 3) lineups. Both target-present and target-absent lineups were created for each of the eight target faces, with foils that were completely independent across the two lineups for each target face.

Eighty participants were recruited to assess the fairness of the 6-person and 12-person lineups via a *mock witness* paradigm (see Malpass & Lindsay, 1999).[1] The proportion of mock witnesses selecting the target (which is compared with the likelihood of identifying the target by chance—i.e., .17 for 6-person lineups and .08 for 12-person lineups) and Tredoux's (1998) *E* (which is compared with the nominal size of the lineup) were used to estimate lineup bias and lineup size, respectively.[2] Across the 6-person lineups, the average proportion of mock witnesses selecting the target face was .18 (95% CI = .15, .20), and the average *E* was 4.99 (95% CI = 4.80, 5.19). Across the 12-person lineups, the average proportion of mock witnesses selecting the target face was .12 (95% CI = .10, .13), and the average *E* was 9.03 (95% CI = 8.80, 9.26).

Procedure

Once the participants were seated in front of the computer, the program began by instructing them that they would be shown a series of faces, which they should study for a subsequent test. Eight faces of White males were displayed sequentially for 3 sec each, with a 1-sec interstimulus interval. The order of presentation of the faces at study was randomized across participants. Following this study phase, the participants completed a distractor task (a digit search puzzle) for 5 min, after which they were prompted by the computer to prepare for a test of their memory.

At test, the participants were instructed that they would be presented with a series of lineup identification tasks in which their goal was to identify the target faces that they had previously viewed. Furthermore, the participants were instructed that within each lineup, *one* of the target faces might or might not be present, and they were provided the options (1) to select a face, (2) to indicate whether a target face was "not present," or (3) to indicate that they were too "unsure" of their memory for the target face to make an identification decision. Following these instructions, the participants were presented with a series of 16 lineup identification tasks, including target-present and target-absent arrays that had been created for each of the eight target faces. Both the presentation of lineups and the placement of faces within each lineup were randomized by the computer program. In addition, the participants were provided with abbreviated instructions prior to each lineup, both as a reminder of their task and as a way to denote the end of one lineup and the beginning of another. For sequential lineups, the presentation of faces ended when the participants identified a face as having been seen previously or

after they had viewed all faces in the lineup without providing an identification response. Following each lineup decision, the participants were asked to provide estimates of confidence ranging from 1 to 7 (*not confident at all* to *very confident*). When the participants had completed the series of 16 lineup identification tasks, they were debriefed by the experimenter and thanked for their assistance.

EXPERIMENT 1
Lineup Presentation and Encoding Strength

As was alluded to previously, one interpretation of the pattern of meta-analytic results observed by Steblay et al. (2001) involves a conservative shift in participants' criterion of responding when they are presented with a sequential lineup (Ebbesen & Flowe, 2002). The first experiment employed our lineup recognition paradigm to examine the effect of lineup presentation method (simultaneous vs. sequential) on identification performance. Given that lineup presentation was predicted to influence the participants' criterion of responding, we also sought a manipulation that might be isolated to discrimination accuracy such that we might further distinguish the effect of the sequential method. Thus, an encoding strength manipulation (see Stretch & Wixted, 1998) was used in which half of our participants were permitted to study the series of faces once and half twice (with each face presented for 3 sec) prior to responding to the lineup identification tasks. We predicted that this encoding strength manipulation would improve discrimination accuracy for those presented with the faces twice, whereas the procedural effects of lineup presentation method would be isolated to participants' criterion of responding (i.e., sequential lineups would heighten response criterion).

Method

Participants. Ninety-six students (58 of them female) from an introduction to psychology course were recruited. Their average age was 21.71 years (*SD* = 4.87).

Design. A 2 × 2 factorial design was used, with lineup presentation (simultaneous vs. sequential) and encoding strength (study once vs. study twice) manipulated between subjects. Twenty-four participants were assigned to each condition.

Procedure. The present experiment followed the general procedures described above, with the exception that the participants shown the faces twice at encoding experienced a 30-sec distractor task (simple math computations) between the two presentations.

Results and Discussion

Mean estimates of the participants' performance are provided in Table 1. Hit and false alarm rates were used to compute SDT estimates of discrimination accuracy (A') and response criterion (B''_D). Both A' and B''_D are nonparametric measures of SDT that Donaldson (1992) has shown to be sufficiently independent (see also Rae, 1976, for a correction to A' when the false alarm rate exceeds the hit rate).

Two 2 × 2 analyses of variance (ANOVAs) were conducted to assess the influence of lineup presentation and encoding strength on each measure. The analysis of the participants' discrimination accuracy scores yielded a significant main effect of the encoding strength manipu-

Table 1
Experiment 1: Identification Performance as a Function
of Lineup Presentation and Encoding Strength

Condition	Hits		False Alarms		A'		B''_D	
	M	SD	M	SD	M	SD	M	SD
Lineup presentation								
Simultaneous	.57	.20	.42	.22	.60	.22	.01	.51
Sequential	.45	.17	.29	.23	.61	.21	.51	.47
Encoding strength								
Study once	.42	.18	.43	.24	.49	.21	.25	.56
Study twice	.60	.17	.28	.21	.72	.16	.26	.55

lation [$F(1,92) = 37.57, MS_e = 0.03, p < .001; \eta^2 = .29$], such that participants who viewed the faces twice showed greater discrimination accuracy than did those who viewed the faces only once. As was expected, there was no significant main effect of the lineup presentation manipulation [$F(1,92) < 1$] and no significant encoding strength $\times$ lineup presentation interaction [$F(1,92) < 1$] on participants' discrimination accuracy scores. In accord with our predictions, the analysis of the participants' response criterion scores demonstrated a significant main effect of lineup presentation [$F(1,92) = 24.57, MS_e = 0.25, p < .001; \eta^2 = .21$], such that sequential lineups (in comparison with simultaneous lineups) engendered more conservative responding. There was no significant main effect of encoding strength [$F(1,92) < 1$] and no significant encoding strength $\times$ lineup presentation interaction [$F(1,92) < 1$] on the participants' response criterion scores.

Taken together, the results of Experiment 1 confirm the criterion shift account of the sequential lineup (Ebbesen & Flowe, 2002), demonstrating that individuals presented with a sequential lineup were significantly more conservative (and therefore less likely to choose) in their identification responses than individuals provided with a simultaneous lineup array. The encoding strength manipulation also had its anticipated effect on individuals' discrimination accuracy, such that a second opportunity for encoding improved subsequent attempts at discrimination. It should also be noted that the effects of these two manipulations were isolated to their respective SDT components, thereby providing some validation for the lineup recognition paradigm introduced in this experiment.

EXPERIMENT 2
Lineup Presentation and Criterion Instructions

Heightening the response criterion of witnesses to a crime at the lineup identification phase provides a reasonable approach to safeguarding against the effects of erroneous identifications. However, this procedure also reduces the likelihood of a correct identification, which law enforcement officials may find compelling enough so as *not* to adopt the method. Some evidence exists (Malpass & Devine, 1981; Steblay, 1997) that a simple criterion-based instruction (i.e., "only respond 'yes' to a

lineup member if you are 100% sure that she or he is the perpetrator that you viewed previously") prior to responding to a lineup will produce a reduction in erroneous identifications but will have no effect on correct identifications. As such, this instructional manipulation would have an influence on both discrimination accuracy and response criterion (i.e., reducing false alarms with no change in hits should lead to a heightened criterion and increased discrimination accuracy). Furthermore, given that a sequential lineup appears to heighten participants' criterion of responding, an instructional manipulation (which has an effect only on false identifications) may provide a useful alternative to altering the method of presentation. In our second experiment, we investigated this possibility within the constraints of the lineup recognition paradigm.

Method

Participants. Ninety-six students (70 of them female) from an introduction to psychology course were recruited. The average age of the sample was 18.94 years ($SD = 1.49$).

Design. A 2 $\times$ 2 factorial design was employed, with lineup presentation (simultaneous vs. sequential) and criterion-based instruction (standard-unbiased vs. conservative-unbiased) manipulated between subjects. Twenty-four participants were tested in each condition.

Procedure. The present experiment followed the general procedures described earlier, with the exception that the participants in the conservative instruction condition were presented with the following general instruction prior to beginning the identification phase of the experiment: "For each lineup that is presented, you should only respond 'yes' that you have seen a face previously if you are completely confident in that decision. If you have any doubts about your memory for the face, please respond 'no' and avoid mistakenly identifying a novel face." In addition, prior to responding to each of the 16 lineups, the participants in the conservative instruction condition were reminded that they should respond "yes" only if they were 100% confident that they had been shown the face previously.

Results and Discussion

Mean estimates of the participants' performance are provided in Table 2. The discrimination accuracy and response criterion scores were examined via 2 $\times$ 2 ANOVAs to assess the influence of lineup presentation and criterion instruction. As was expected, the analysis of participants' discrimination accuracy scores yielded a significant main effect of the criterion instruction manipulation [$F(1,92) = 4.95, MS_e = 0.04, p < .05; \eta^2 = .05$]. Par-

Table 2
Experiment 2: Identification Performance as a Function of Lineup
Presentation and Criterion Instructions

Condition	Hits		False Alarms		A'		B''_D	
	M	SD	M	SD	M	SD	M	SD
Lineup presentation								
Simultaneous	.48	.19	.40	.20	.56	.19	.18	.49
Sequential	.38	.19	.30	.19	.57	.21	.57	.39
Criterion instruction								
Standard	.45	.19	.42	.19	.53	.21	.27	.46
Conservative	.42	.20	.28	.19	.62	.18	.48	.48

ticipants who had been instructed to heighten their criterion showed some improvement in their discrimination accuracy, when compared with participants in the standard instruction condition. As in Experiment 1, there was no significant main effect of lineup presentation on participants' discrimination accuracy [$F(1,92) < 1$] and no significant criterion instruction × lineup presentation interaction [$F(1,92) < 1$]. In accord with our predictions, the analysis of participants' response criterion scores yielded a significant main effect of the lineup presentation manipulation [$F(1,92) = 19.70, MS_e = 0.19, p < .001; \eta^2 = .18$], such that sequential lineups (in comparison with simultaneous lineups) once again engendered more conservative responding. In addition, a significant main effect of the criterion instruction manipulation was present [$F(1,92) = 5.49, MS_e = 0.19, p < .05; \eta^2 = .06$]. Those provided with conservative instructions were significantly more conservative in their identification responses, when compared with those provided with standard instructions. The criterion instruction × lineup presentation interaction was not significant [$F(1,92) < 1$].

In accord with predictions, sequential lineups once again prompted a conservative shift in the participants' response criterion but had no effect on overall discrimination accuracy. In contrast, instructions to respond to lineups with a conservative criterion (i.e., "only respond to a lineup member if you are 100% confident") led to both a significant improvement in discrimination accuracy and a conservative shift in criterion of responding. This dual effect of criterion instructions was largely the result of a significant decrease in false alarms for the conservative instruction condition [$F(1,92) = 11.86, MS_e = 0.04, p < .001; \eta^2 = .11$], with no concomitant reduction in hits [$F(1,92) < 1$], a finding that is consistent with previous studies in the eyewitness identification literature (see Steblay, 1997).

One might reasonably inquire how the two criterion manipulations in the present experiment (both occurring at the time of test) could demonstrate such differential effects on participants' performance and, in particular, why an instruction-based criterion manipulation would also influence discrimination accuracy. One interpretation of the instruction effect is to infer that the change in discrimination may have occurred as a by-product of a weak criterion manipulation. Such an explanation could be accounted for by first assuming that the underlying familiarity distribution for the sample of lure faces (lineup foils) had a greater variance than did the distribution of target faces—a reasonable assumption, given that the ratio of lures to targets was 6:1—and second by recognizing that the shift in criterion was greater for the lineup presentation manipulation ($\eta^2 = .18$) than for the criterion instruction manipulation ($\eta^2 = .06$). Given this set of circumstances, the change in discrimination accuracy observed with the criterion instruction could be explained as the by-product of a weak shift in the criterion not causing a change in the hit rate. In contrast, the more substantial shift in criterion for the lineup presentation manipulation led to a significant drop in both hit and false alarm responses, thereby not influencing participants' discrimination accuracy scores.

EXPERIMENT 3
Lineup Presentation and Nominal Lineup Size

The Technical Working Group for Eyewitness Evidence (1999) has recommended that a minimum of five foils (or nonsuspects) be included in any photograph lineup. The majority of U.S. law enforcement agencies follow this recommendation by employing photograph lineups containing 6 to 8 individuals, whereas the British and Canadian systems generally construct lineups with 9 or 12 individuals, respectively. Although several studies using the classic eyewitness identification paradigm have demonstrated little or no effect of increasing lineup size (Cutler, Penrod, & Martens, 1987; Cutler, Penrod, O'Rourke, & Martens, 1986; Nosworthy & Lindsay, 1990), Levi and Lindsay (2001) have proposed that increasing the size of lineups should theoretically reduce the likelihood of false identifications of innocent suspects but have little or no influence on correct identifications of the guilty. From its theoretical perspective, SDT actually predicts that participants' criterion of responding should become more conservative as lineup size is increased, thus involving a reduction in both hits and false alarms. For example, Xu and Bellezza (2001) have shown that increasing the number of foils in a recognition task leads participants to heighten their criterion of responding (i.e., via a reduction in both hits and false alarms) as a result of their perceptions regarding the likelihood of iden-

tifying a target stimulus. Theoretically, such an effect should hold for lineup arrays that vary in size, as well as show-ups (or single-target identifications), particularly when chance responding as a function of lineup size is taken into consideration (see Steblay, Dysart, Fulero, & Lindsay, 2003).

Method

Participants. Two hundred sixty students (179 of them female) from an introduction to psychology course were recruited. Their average age was 19.67 years ($SD = 3.06$).

Design. A 2×6 factorial design was employed, with lineup presentation (simultaneous vs. sequential) and nominal lineup size (2-, 4-, 6-, 8-, 10-, or 12-person arrays) manipulated between subjects. An additional *show-up* condition was conducted for comparison purposes. Twenty participants were tested in each condition. Once again, the participants responded to both target-present and target-absent lineups, thereby providing two measures (hits and false alarms) of identification performance. These measures were subsequently used to compute SDT measures of discrimination accuracy (A') and response criterion (B''_D).

Procedure. The present experiment followed the general procedures described earlier, with the exception that nominal lineup size was varied across participants. PC_Eyewitness was adjusted such that lineup foils were selected randomly from among 12 possible foils for each lineup, except that all the foils were presented when the nominal size was 12.

Results and Discussion

Mean estimates of the participants' performance are provided in Table 3. Given that predictions are based on the consideration of chance responding across lineup size, the participants' false alarm responses were corrected on the basis of the nominal size of the lineup array (i.e., proportion of false alarms divided by nominal size of the lineup), and these corrected estimates were used to compute SDT parameters.

Two 2×6 ANOVAs were conducted to assess the influence of lineup presentation and nominal lineup size on the participants' discrimination accuracy and response criterion scores. Results indicated a significant main effect of lineup size on discrimination accuracy [$F(5,247) = $ 2.66, $MS_e = 0.01$, $p < .05$; $\eta^2 = .05$], such that increasing lineup size caused some reduction in accuracy of discrimination. Neither the main effect of lineup presentation [$F(1,247) = 2.30$] nor the lineup size $\times$ lineup presentation interaction [$F(5,247) < 1$] was significant. In accord with our predictions, analysis of the participants' response criterion scores demonstrated a significant main effect of lineup presentation [$F(1,247) = 21.70$, $MS_e = 0.06$, $p < .001$; $\eta^2 = .08$], such that sequential lineups once again engendered more conservative responding. In addition, a rather large effect of lineup size was observed [$F(5,247) = 38.97$, $MS_e = 0.06$, $p < .001$; $\eta^2 = .44$], demonstrating that increasing lineup size led to more conservative responding. Finally, the lineup size $\times$ lineup presentation interaction was also significant [$F(5,247) = 3.31$, $MS_e = 0.06$, $p < .01$; $\eta^2 = .06$]. As is displayed in Figure 1, follow-up analyses indicated that sequential lineups engendered more conservative responding than did simultaneous lineups [ts(38) ≥ 2.06, ps $< .05$], except when nominal size reached 12 persons [$t(38) < 1.00$]. With regard to the standard (uncorrected) estimates, the pattern of results was quite consistent with that of the corrected estimates, with the exception that the lineup presentation $\times$ nominal lineup size interaction was not significant [$F(5,228) < 1$].

Multiple planned comparisons were also conducted (with a corrected significance level of $p < .01$) to assess the performance of participants presented with show-ups versus lineups of increasing nominal size (2, 4, 6, 8, 10, or 12). Similar to the main effect of nominal lineup size, show-ups engendered a significantly more liberal criterion of responding when compared with nominal lineup sizes greater than 4 persons [ts(58) > 4.07, ps $< .001$], while also demonstrating a marginal improvement in discrimination accuracy when compared with the same conditions [ts(58) > 2.82, ps $< .01$].

In accord with our previous experiments and SDT predictions, sequential lineups led to a conservative shift in the criterion of the participants, when compared with performance on simultaneous arrays. The present exper-

Table 3
Experiment 3: Identification Performance as a Function of Lineup Presentation and Nominal Lineup Size

| | Hits | | False Alarms | | | | A' | | | | B''_D | | | |
| | | | Standard | | Corrected | | Standard | | Corrected | | Standard | | Corrected | |
Condition	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Lineup presentation														
Simultaneous	.56	.23	.39	.21	.07	.05	.62	.26	.84	.10	.09	.54	.70	.38
Sequential	.46	.21	.31	.21	.05	.06	.61	.23	.82	.08	.42	.45	.85	.22
Nominal lineup size														
Show-up	.80	.22	.08	.08			.91	.11			.34	.45		
2-person array	.73	.17	.28	.19	.14	.10	.79	.18	.87	.10	−.05	.59	.33	.45
4-person array	.59	.16	.34	.19	.09	.05	.68	.15	.85	.06	.16	.50	.71	.26
6-person array	.51	.21	.38	.23	.06	.04	.59	.25	.82	.10	.23	.50	.85	.14
8-person array	.45	.22	.37	.21	.05	.03	.55	.24	.82	.09	.34	.47	.89	.17
10-person array	.40	.19	.38	.21	.04	.02	.53	.23	.81	.07	.38	.44	.94	.07
12-person array	.38	.20	.34	.22	.03	.02	.56	.30	.80	.12	.47	.48	.95	.07

Note—Standard cells are computed on the basis of the proportion of false identifications regardless of lineup size. Corrected cells are based on the proportion of false identifications divided by the nominal size of the lineup array.
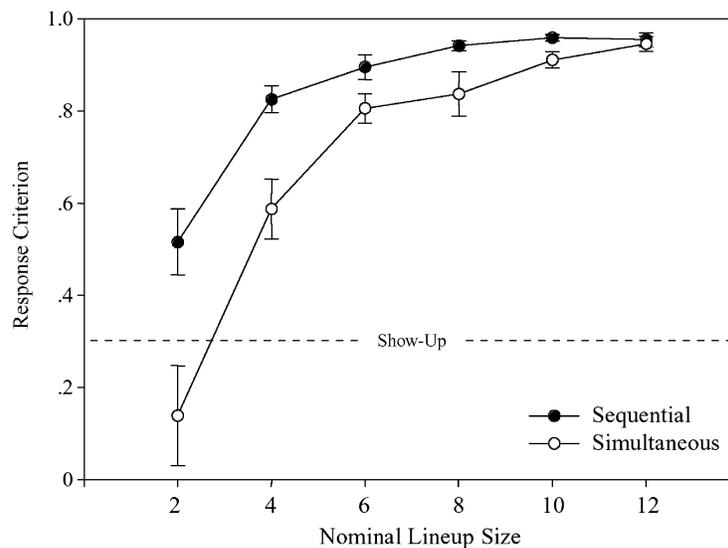
Figure 1. Response criterion of the participants in simultaneous and sequential lineups as a function of nominal lineup size. Error bars represent standard error estimates above and below each mean.

iment also demonstrated that increasing lineup size led to a conservative shift in the participants' criterion of responding. This heightened criterion was likely the result of a shift in participants' estimations of the likelihood of identifying a target stimulus, given the increased frequency of novel stimuli (Xu & Bellezza, 2001). Finally, these two manipulations (i.e., lineup presentation and lineup size) also interacted in such a way that any differences in the participants' response criterion when they responded to simultaneous versus sequential lineups were statistically eliminated when lineup size was 12.

EXPERIMENT 4
Dual-Process Estimates
Across Lineup Presentation

Dual-process theories of memory posit a distinction between two independent cognitive operations—namely, an episodic-based recollection process and a fluency-based familiarity process (Gardiner & Richardson-Klavehn, 2000; Kelley & Jacoby, 2000). Studies in which these two processes have been investigated have demonstrated their unique influence on memory through dissociations in which a given manipulation is shown to influence one process or the other. For example, in his review of the literature, Yonelinas (2002) demonstrated that encoding-based manipulations will generally influence *recollection*, whereas *familiarity* is often sensitive to fluency manipulations, forgetting over short retention intervals, and changes in response criterion.

With regard to the conservative criterion shift demonstrated by sequential lineups in the three previous experiments, a dual-process framework would seem to predict a reduction in familiarity-based responding when stim-

uli are presented sequentially. More specifically, perceived familiarity may provide a stronger basis for memory decisions in simultaneous arrays, given that presenting stimuli at the same time can provide a contextual basis for the application of familiarity (possibly consistent with the *relative judgment* decision strategy proposed by Lindsay & Wells, 1985). In contrast, reliance upon familiarity-based processes may be reduced when faces are presented independently from one another, since the contextual basis that facilitates such judgments is eliminated. An alternative prediction has recently been provided by Gronlund (2005), who proposed that the sequential advantage may be largely due to the application of recollection-based processes, particularly when a stimulus is encoded in a "distinctive" manner. This final experiment provides a basis for estimating the contributions of recollection and familiarity in simultaneous versus sequential lineups (and thereby distinguishing these two predictions) by asking participants to provide remember–know–guess judgments following each identification decision.

Method

Participants. Eighty students (52 of them female) from an introduction to psychology course were recruited. The average age of the sample was 23.05 years ($SD = 5.38$).

Design. A single-factor design was employed, with lineup presentation (simultaneous vs. sequential) manipulated between subjects. Forty participants were tested in each condition. As in the previous experiments, the participants responded to both target-present and target-absent lineups, thereby providing two measures (hits and false alarms) of identification performance. These measures were used to compute SDT measures of discrimination accuracy (A') and response criterion (B''_D).

Procedure. The present experiment followed the general procedures described earlier, with the exception that the participants were

asked to provide a remember–know–guess judgment following each identification decision (see Gardiner & Richardson-Klavehn, 2000).

Results and Discussion

Mean estimates of the participants' identification performance and remember–know judgments are provided in Table 4. In accord with the previous experiments, sequential lineups led to a significant conservative shift in the participants' response criterion [$t(78) = 4.93$, $p < .001$] but produced no significant change in discrimination accuracy scores [$t(78) < 1$]. Given the general assumptions of independence governing the dual-process model (see Kelley & Jacoby, 2000), estimates of know responses were corrected for independence for each participant, using the *independence remember–know* procedure [i.e., $K_C = K/(1-R)$; see Yonelinas & Jacoby, 1995] prior to analyzing differences in responding to simultaneous versus sequential lineups. As is predicted by a dual-process framework (Yonelinas, 2002), sequential lineups led to a significant reduction in familiarity-based responding [$t(78) = 2.76$, $p < .01$] but had no significant effect on recollection-based judgments [$t(78) = 1.55$, $p = .13$]. When the results from Experiment 4 are considered as a whole, it appears that the relative similarity between lineup members may create a context in which fluency-based processing is facilitated to a greater extent when lineup members are presented simultaneously but is reduced when lineup members are presented sequentially. In contrast to Gronlund's (2005) suggestion, changes in lineup presentation did not alter the frequency of remember judgments in the present experiment.

GENERAL DISCUSSION

The sequential lineup has been a much heralded innovation in eyewitness identification methods; indeed, it has been argued that law enforcement should adopt such a procedure for everyday investigative practice (Levi & Lindsay, 2001; Lindsay, 1999). However, we are just beginning to understand the cognitive processes governing such decisions. In the present experiments, we sought to empirically verify the criterion shift explanation for sequential lineup decisions offered by Ebbesen and Flowe (2002) and to further extend such an analysis by estimating the contributions of recollection and familiarity processes to memory performance. To accomplish such analyses, it was necessary to introduce a novel lineup recognition paradigm in which participants were presented with multiple target faces at encoding and were

later asked to identify these target faces from a series of target-present and target-absent lineup arrays. The paradigm permitted us to gather multiple observations from each participant and, thereby, compute SDT estimates of discrimination accuracy and response criterion. Across four experiments, our results consistently demonstrated a conservative shift in response criterion as the participants were exposed to simultaneous versus sequential lineup arrays.

Dual-process theories (for a review, see Yonelinas, 2002) hold that two independent parameters contribute to memory performance—namely, recollection and familiarity. Previous studies in this literature have indicated that shifts in the response criterion are frequently associated with fluctuations in familiarity processes; however, shifts in criterion rarely influence the use of recollection in memory decisions (Yonelinas, 2002). As such, it was expected that responses to sequential lineups would be less influenced by familiarity-based processes, when compared with responses to simultaneous lineups. In contrast, Gronlund (2005) has suggested that the sequential advantage may lie in greater reliance upon recollection-based processes. Using the remember–know–guess procedure (see Gardiner & Richardson-Klavehn, 2000), we examined the contributions of recollection and familiarity across simultaneous versus sequential lineups. In accord with previous studies in the dual-process literature, it appears that the participants were more likely to use familiarity-based processes when responding to simultaneous, as opposed to sequential, lineups.

Such a finding may be quite consistent with the notion of *relative* judgments alluded to by Lindsay and Wells (1985), since the relative similarity between lineup members may create a context in which fluency-based processing is facilitated to a greater extent than when lineup members are presented independently. Although familiarity-based responding can improve target-present identification performance when retrieval conditions are sufficiently heterogeneous (e.g., sufficient differences in perceived familiarity between the target and the lineup foils are present), reliance upon familiarity can also hinder performance in the target-absent condition in which a biasing factor is present (e.g., use of a high-similarity "innocent suspect" or clothing that is similar to that worn by the target person). Thus, promoting reliance upon familiarity-based responding may be thought of as a two-edged sword.

Although simultaneous arrays appear to promote familiarity-based processing, we are somewhat hesitant

Table 4
Experiment 4: Identification Performance and Frequency of Remember–Know Judgments
as a Function of Lineup Presentation

Lineup Presentation	Hits		False Alarms		A'		B''_D		R		K_C	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Simultaneous	.56	.15	.56	.21	.49	.20	−.22	.48	.24	.17	.41	.21
Sequential	.37	.23	.45	.23	.45	.26	.33	.52	.30	.22	.28	.22

Note—R, remember judgments; K_C, know judgments corrected for independence.

to suggest that sequential arrays could provide a practical alternative, given the conservative shift in response criterion that occurs. Rather, a more diagnostic alternative would involve procedures that isolate the change in performance to *false identifications*. One procedure that appears to meet this criterion in the present experiments involved instructing the participants to respond "yes" to a target face only if they were 100% confident that they had seen the face previously and to avoid mistakenly identifying a face that had not been presented. This manipulation influenced decisions across both simultaneous and sequential arrays, improving diagnosticity by approximately 50% (i.e., diagnosticity, as measured by proportion of correct identifications divided by the proportion false identifications, was 1.07 for the standard instruction condition vs. 1.50 for the conservative instruction condition).

Finally, we wish to note the utility of the lineup recognition framework introduced in the present studies for understanding eyewitness memory. Although we have focused here on various procedural and structural aspects of lineup identification, we believe that this paradigm will prove useful to researchers investigating other eyewitness memory factors. For example, researchers could use the lineup recognition framework to examine the variety of estimator and system variables discussed by Wells and Olson (2003), such as the cross-race effect (Meissner & Brigham, 2001) or foil selection strategies (Luus & Wells, 1991), respectively. Furthermore, SDT and dual-process memory approaches provide a firm theoretical basis from which to understand the cognitive and social psychological mechanisms underlying a phenomenon and from which predictions can be generated in attempts to improve the diagnostic value of eyewitness identification evidence.

REFERENCES

Brigham, J. C., Meissner, C. A., & Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology*, **13**, S73-S92.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology*, **72**, 629-637.

Cutler, B. L., Penrod, S. D., O'Rourke, T. E., & Martens, T. K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour*, **1**, 113-134.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, **121**, 275-277.

Ebbesen, E. B., & Flowe, H. D. (2002). *Simultaneous v. sequential lineups: What do we really know?* Retrieved December 20, 2001, from http://www.psy.ucsd.edu/~eebbesen/SimSeq.htm.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229-244). London: Oxford University Press.

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

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

Huff, C. R., Rattner, A., & Sagarin, E. (1996). *Convicted but innocent: Wrongful conviction and public policy*. Thousand Oaks, CA: Sage.

Kelley, C. M., & Jacoby, L. L. (2000). Recollection and familiarity: Process-dissociation. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 215-228). London: Oxford University Press.

Levi, A. M., & Lindsay, R. C. L. (2001). Lineup and photo spread procedures: Issues concerning policy recommendations. *Psychology, Public Policy, & Law*, **7**, 776-790.

Lindsay, R. C. L. (1999). Applying applied research: Selling the sequential line-up. *Applied Cognitive Psychology*, **13**, 219-225.

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

Luus, C. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law & Human Behavior*, **15**, 43-57.

MacLin, O. H., Meissner, C. A., & Zimmerman, L. A. (2005). PC_Eyewitness: A computerized framework for the administration and practical application of research in eyewitness psychology. *Behavior Research Methods*, **37**, 324-334.

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

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

Malpass, R. S., & Lindsay, R. C. L. (1999). Measuring line-up fairness. *Applied Cognitive Psychology*, **13**, S1-S7.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law*, **7**, 3-35.

Nosworthy, G. J., & Lindsay, R. C. L. (1990). Does nominal lineup size matter? *Journal of Applied Psychology*, **75**, 358-361.

Rae, G. (1976). A non-parametric measure of recognition performance. *Perceptual & Motor Skills*, **42**, 98.

Scheck, B. C., & Neufeld, P. J. (2001). DNA and innocence scholarship. In S. Westervelt & J. Humphreys (Eds.), *Wrongly convicted: Perspectives on failed justice* (pp. 241-252). New Brunswick, NJ: Rutgers University Press.

Scheck, B. C., & Neufeld, P. J., & Dwyer, J. (2000). *Actual innocence: Five days to execution and other dispatches from the wrongly convicted*. New York: Doubleday.

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

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law & Human Behavior*, **25**, 459-474.

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

Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1379-1396.

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

Tredoux, C. G. (1998). Statistical inference on lineup measures. *Law & Human Behavior*, **22**, 217-237.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, **54**, 277-295.

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, & Cognition*, **27**, 1197-1210.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A

review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517.

YONELINAS, A. P., & JACOBY, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory & Language*, **34**, 622-643.

NOTES

1. Mock witness evaluations are widely used in eyewitness research to test the fairness of lineups (see Brigham, Meissner, & Wasserman, 1999; Malpass & Lindsay, 1999). A group of mock witnesses, who have not witnessed the crime event in question and who are blind to the identity of the perpetrator, are typically provided a brief description of the perpetrator and are asked to select the suspect from the lineup on the basis of this description. If the lineup is fair, the mock witnesses should not be able to identify the suspect at a rate greater than chance (referred to as *lineup bias*), and the distribution of their choices should be spread equally over the lineup members (referred to as *lineup size*).

2. Tredoux's (1998) E is a measure that estimates the number of plausible lineup members by subtracting a quantity from the number of nominal lineup members. The quantity is determined as a function of the deviation of mock witness choices from uniformity.