Journal of Experimental Psychology: Applied
2010, Vol. 16, No. 4, 387–398

1076-898X/10/$12.00 DOI: 10.1037/a0021034

Postidentification Feedback Affects Subsequent Eyewitness Identification Performance

Matthew A. Palmer, Neil Brewer, and Nathan Weber
Flinders University

Eyewitnesses sometimes view more than one lineup during an investigation. We investigated the effects of postidentification feedback following one lineup on responses to a second lineup. Witnesses ($N = 621$) viewed a mock crime and, later, attempted to identify the culprit from an initial (target-absent) lineup and a second (target-present or target-absent) lineup. Prior to viewing the second lineup, some witnesses received accurate feedback stating that the initial lineup did not contain the culprit. A compound-decision, signal detection approach allowed the effects of feedback on identification responses to be described in terms of differences in discriminability and response bias. For witnesses who made an incorrect foil identification from the initial lineup, feedback (vs. no feedback) was associated with poorer discriminability on the second test. For witnesses who correctly rejected the initial lineup, feedback (vs. no feedback) was associated with greater discriminability on the second test. Only witnesses who received feedback after an initial correct rejection performed at a level comparable with a single-lineup control group, suggesting that an initial identification test can impair, but not enhance, performance on a second test involving the same culprit. From a theoretical perspective, the results are consistent with the idea that the way people use memorial information when making memory decisions is flexible. Analyses of preidentification confidence ratings, obtained in a follow-up study ($N = 133$), suggested that the effects of feedback on identification performance may have operated via differences in witnesses' metacognitive beliefs.

Keywords: eyewitness identification, metacognition, multiple lineups, postidentification feedback, signal detection theory

Laboratory research (e.g., Steblay, Dysart, Fulero, & Lindsay, 2001) and field studies (Cutler & Penrod, 1995) have shown that eyewitness identification errors are not uncommon, and false identifications have been cited as the most common cause of wrongful conviction in Great Britain and North America (Innocence Project, 2009; Smith, Lindsay, & Pryke, 2000). It is not surprising, then, that much research has investigated factors that affect the likely accuracy of identification decisions. One variable that has received attention is the use of multiple identification procedures (e.g., Deffenbacher, Bornstein, & Penrod, 2006; Godfrey & Clark, 2010). In this research, we examined a novel issue: whether feedback following an initial identification decision influences performance on a second identification test. There is evidence that feedback can influence witnesses' source memory for details of a mock crime (Lane, Roussel, Villa, & Morita, 2007), and the effects of postidentification feedback on confidence judgments have been well-documented (e.g., Bradfield, Wells, & Olson, 2002; Semmler & Brewer, 2006; Semmler, Brewer, & Wells, 2004; Wells & Bradfield, 1998, 1999). However, the effects of postidentification feedback on responses to a subsequent identification test have not been studied.

Multiple Identification Tests and the Scope of the Present Study

Prior to viewing a lineup, a witness may complete an initial identification procedure. The initial identification procedure can involve a show-up (i.e., where a single person is presented to the witness; e.g., Godfrey & Clark, 2010), a mugshot search (see Deffenbacher et al., 2006, for a review), or another lineup (Hinz & Pezdek, 2001; Pezdek & Blandon-Gitlin, 2005; Wells, 1984). Archival data suggest that it is not uncommon for criminal investigations to involve the use of multiple identification procedures (Behrman & Davey, 2001), including multiple lineups. For example, in one set of data from cases in the U.K., 129 (15.2%) of 851 eyewitnesses viewed two or more lineups during the course of police investigations (Halford, 2009).

Because scenarios involving multiple identification procedures can vary considerably (e.g., in terms of the type of initial test used), it is important to specify the scope of the present investigation. First, this research involves an initial lineup, rather than a showup or mugshot search. Second, we focused on the simplest multiple-lineup scenario, whereby witnesses view two lineups, both of which were designed to ascertain the identity of the same culprit. This scenario can arise, for example, when the crime in question involves a single culprit but the police have more than

Matthew A. Palmer, Neil Brewer, and Nathan Weber, School of Psychology, Flinders University.

This research was supported by Australian Research Council Grants DP0556876 and DP1093210. We gratefully acknowledge Paul Williamson and James Sauer for providing many insightful comments, and Nicole Reid and Emily Palmer for assistance with data collection.

Correspondence concerning this article should be addressed to Matthew A. Palmer, School of Psychology, Flinders University, GPO Box 2100, Adelaide, South Australia, Australia. E-mail: matthew.palmer@flinders.edu.au

one suspect. In such cases, the police may construct a separate lineup for each suspect. This scenario can also occur if the police initially arrest one suspect and have the witness attempt to identify him or her from a lineup; later, new evidence may come to light exonerating the initial suspect and implicating another who may, in turn, be placed in a second lineup for the witness to view. Third, unlike the vast majority of studies involving multiple identification procedures, the present study does not use a repeated-foil design (i.e., where one or more innocent persons viewed in the initial identification procedure also appeared in a subsequent lineup). Although repeated-foil designs have considerable applied value in the context of mugshot studies (e.g., Deffenbacher et al., 2006), Pezdek and Blandon-Gitlin (2005) note that such designs are of limited relevance to multiple-lineup scenarios, since it is highly unlikely that an innocent person would appear in more than one lineup during the course of an appropriately conducted investigation. In the present research, no person (foil or culprit) appeared in more than one lineup. In sum, this research examined the effects of feedback following an initial lineup response on witnesses' responses to a second lineup (not including any members of the initial lineup) for the same culprit.

Postidentification Feedback

During criminal investigations, eyewitnesses sometimes receive feedback following an identification decision (Nettles, Nettles, & Wells, 1996). Postidentification feedback can be confirming (e.g., "Good. You identified the suspect") or disconfirming (e.g., "Actually, the person you chose was not the suspect"). In the present study, we examined two types of feedback that might plausibly be given to a witness in a multiple-lineup scenario involving a single culprit: confirming feedback following a lineup rejection (e.g., "you were right to reject the earlier lineup as we now have evidence proving that suspect innocent"), and disconfirming feedback following a positive identification (e.g., "the person you chose from the last lineup was actually someone we put in there who we know is not the culprit"). Such feedback provides witnesses with two types of information that might influence responses to a second lineup.

Effects of feedback on response bias. First, postidentification feedback provides witnesses with information about the likelihood that the culprit will appear in a second lineup. Witnesses who choose from an initial lineup (henceforth termed "initial choosers") may assume that the culprit, having presumably appeared in one lineup, will not appear in a second lineup (cf. Davies, Shepherd, & Ellis, 1979). Feedback implying that the culprit was not present in the initial lineup should undermine this assumption, resulting in a lenient shift in response bias (i.e., the tendency to choose from or reject the second lineup) relative to no feedback. Note that we use the term "decision criterion" to refer to the amount of evidence required for a positive response to be made, and the term "response bias" to describe the tendency to respond positively or negatively, relative to optimal (unbiased) responding. Thus, a lenient shift in response bias refers to a lenient shift in the decision criterion relative to the point of unbiased responding.

In contrast to initial choosers, witnesses who reject an initial lineup (henceforth, "initial nonchoosers") would not be expected to make any assumptions about the likely presence of the culprit in a second lineup. Thus, feedback confirming that the culprit was not present in an initial lineup should not affect response bias for initial nonchoosers. (For further discussion of the effects of target base-rate information on criterion placement, see Hirshman, 1995; Rotello, Macmillan, Reeder, & Wong, 2005; Strack & Forster, 1995.)

The second type of information conveyed by postidentification feedback relates to witnesses' metacognitive beliefs. (We use the terms metacognitive and metamnemonic interchangeably.) In the simplest multiple-lineup scenario (tested here), a witness would receive feedback for only one prior identification response before viewing a second lineup. This feedback would be entirely confirming or disconfirming. Success feedback has been shown to have a positive influence (and failure feedback a negative influence) on metacognitive variables such as perceptions of ability and expectations of success on a variety of cognitive tasks (e.g., Anderson, White, & Wash, 1966; Bouffard-Bouchard, 1993; Bridgeman, 1974; Feather, 1966; Feather & Saville, 1967; Osler, 1954) including recall memory tests (e.g., Klein, Loftus, & Fricker, 1994). To the extent that these findings extend to eyewitness identification decisions, confirming feedback should have a positive (and disconfirming feedback a negative) effect on witnesses' perceptions of, for example, their ability to perform identification tests involving the culprit, or the quality of their memory for the culprit. Consistent with this idea, Wells and Bradfield (1998) found that, compared to a no-feedback control condition, confirming feedback increased (and disconfirming feedback decreased) witnesses' ratings of the amount of attention paid to the witnessed event, the ease with which identification decisions were made, the extent to which the identification decision was based on sufficient information, and willingness to testify in court.

Such effects of feedback on witnesses' metacognitions might be expected to translate into effects on response bias if the witness attempts a second identification test. Stretch and Wixted (1998, Experiment 1) found that participants adjusted their decision criterion in response to between-list changes in the strength of target items on a word recognition test. Presumably, participants intuitively attempted to maximize accuracy by adopting a more conservative criterion for strong items and a more lenient criterion for weak items. (For a more detailed discussion of strength-based criterion shifts, see Stretch & Wixted). In a similar vein, witnesses who believe that their memory for a culprit is poor (e.g., because they received disconfirming feedback stating that they responded incorrectly on a previous identification test for that person) might be expected to set a more lenient decision criterion on a second test. Conversely, witnesses who believe that their memory for a culprit is good (e.g., after receiving confirming feedback for a previous test) might be expected to set a more conservative criterion. Because such criterion shifts would not be accompanied by actual differences in memory strength, the net effect of disconfirming feedback would be a lenient shift in response bias and the net effect of confirming feedback would be a conservative shift in bias.

Effects of feedback on discriminability. Although the idea is perhaps unlikely to have received much thought from eyewitness identification or recognition memory researchers, there is some reason to think that feedback might influence discriminability (i.e., witnesses' ability to distinguish the culprit from foils) rather than response bias on a second identification test. We propose one possible mechanism by which feedback might pro-

duce greater discriminability following confirming feedback (vs. no feedback) and poorer discriminability following disconfirming feedback (vs. no feedback). This notion is based on several theoretical approaches, including the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993) and multidimensional models of memory (e.g., Banks, 2000), which hold that the way people use memorial information when making memory decisions is malleable. (We use the terms memorial information and memorial evidence interchangeably.) For instance, differences in metacognitive beliefs can produce differences in the way that memorial information is used and, hence, differences in memory performance (e.g., Jacoby, Shimizu, Daniels, & Rhodes, 2005; Johnson et al., 1993; Lane et al., 2007; Starns, Lane, Alonzo, & Roussel, 2007).

Of particular relevance to the present research, there is evidence that the perceived strength of memories can influence rememberers' expectations about the amount and type of information that will be available during memory decisions. For example, Lloyd, Westerman, and Miller (2003, Experiment 3) found that a manipulation of fluency (via priming) had less effect on test responses when participants believed that they had studied items five times rather than once. (In fact, participants had studied a counterfeit list which contained no target items.) Lloyd et al. concluded that people expect greater processing fluency when recognizing items they believe are encoded in memory relatively strongly. Such differences in expectations can result in memorial information being evaluated more, or less, stringently during memory decisions (Johnson et al., 1993). For example, Schacter and colleagues found lower false alarm rates for items encoded as pictures rather than words (Schacter, Israel, & Racine, 1999; Dodson & Schacter, 2002) and items that were said rather than heard (Dodson & Schacter, 2001), and concluded that people expect more vivid, distinctive information to be available when recognizing pictures (vs. words) and said items (vs. heard items). It is important that there is evidence that these effects are not simply due to familiarity-based criterion shifts, and do not rely on differences in the actual strength of pictures versus words or said versus heard items (Gallo, Weiss, & Schacter, 2004). Thus, when participants evaluated memorial evidence more stringently by placing greater weight on the presence of distinctive information, they were able to better distinguish between studied and unstudied items at test.

In a similar vein, we suggest that witnesses who believe that their memory for a culprit is relatively strong (e.g., after confirming feedback vs. no feedback) will evaluate memorial evidence more stringently during a second identification decision. In turn, discriminability on the second test will be better following confirming feedback versus no feedback. Conversely, witnesses who believe that their memory for a culprit is relatively weak (e.g., after disconfirming feedback vs. no feedback) will evaluate memorial evidence less stringently during a second identification decision. In turn, discriminability on the second test will be poorer following disconfirming feedback versus no feedback. Note that we do not make any specific claims about the way that feedback might influence participants' use of different types of memorial information (e.g., whether feedback causes people to weight some types of information more or less heavily, to ignore some types of information entirely, or to take new information into account). Instead, following Lane et al., (2007), we assume that the types of memorial information that produce better or worse memory performance can vary between people. In turn, what might constitute more or less stringent evaluation of memorial evidence can also vary between people.

Estimating Discriminability and Response Bias for Compound Decisions

Testing the hypotheses described earlier requires the ability to disentangle differences in discriminability from differences in response bias (i.e., the tendency of witnesses to choose from or reject a lineup). This is typically achieved by calculating numerical indices of discriminability (e.g., d') and bias (e.g., c). Calculation of these indices is straightforward for tasks involving simple decisions (e.g., old/new recognition tests). However, an eyewitness identification decision is, in signal detection terms, a compound decision, comprising *detection* and *identification* components (Macmillan & Creelman, 1991).

The detection component refers to the ability to detect the presence of a target stimulus. An eyewitness identification test includes a 1-of-m detection component, referring to witnesses' ability to detect the presence of the target in a lineup comprising m stimuli. For the detection component of an identification decision, differences in choosing rates that are uniform across target-present and target-absent lineups reflect differences in response bias (e.g., higher choosing from target-present and target-absent lineups indicates more lenient bias). However, differences in choosing rates that are moderated by target-presence reflect differences in detection discriminability. For example, greater discriminability is evidenced by lower choosing from target-absent lineups and higher choosing from target-present lineups (regardless of whether the target or a foil is chosen).

The identification component (termed SDT-identification henceforth to distinguish it from the more general term "identification") refers to the ability of a perceiver to attribute the correct identity to a stimulus or set of stimuli. An eyewitness identification test includes an m-alternative, forced-choice SDT-identification component. This reflects witnesses' ability to correctly identify the target, given that a positive response has been made to a target-present lineup. For the SDT-identification component of an eyewitness decision, greater discriminability is evidenced by a greater ratio of correct identifications to foil identifications from target-present lineups. Response bias is not relevant for the SDT-identification component, as a positive identification is assumed.

Only four possible responses can be made in simple decisions: hits and misses for target-present trials, and false alarms and correct rejections for target-absent trials. In contrast, the compound nature of eyewitness identification decisions results in an extra type of response: foil identifications from target-present lineups. These responses are problematic for the calculation of d' and c, as they could be classified as both false alarms (as a foil was chosen) and misses (as the target was present but not chosen). Due to a lack of models that deal with this problem (Duncan, 2006, 2010), eyewitness studies that have examined discriminability have had to rely on proxy estimates, such as the difference between hits and false alarms (Gronlund, 2004) or d' calculated with foil identifications from target-present lineups excluded (Meissner, Tredoux, Parker, & MacLin, 2005). In the present study, SDT-Compound Decision (SDT-CD; Duncan, 2006, 2010; see also Starr, Metz, Lusted, &

Goodenough, 1975), a model designed specifically for compound decisions, was used to generate estimates of d' and c. This enables the effects of feedback to be described in terms of differences in estimates of discriminability and bias. In the model, expected probabilities of false alarms (i.e., positive responses to target-absent trials), hits (i.e., positive responses to target-present trials), and correct identifications are calculated for all possible combinations of a single d' value and a single decision criterion value. (The c estimate is derived from the d' and criterion values, with c = criterion – $d'/2$.) The model-generated probabilities are then compared to observed response probabilities, and the combination of parameters that best fit the observed data for each condition is selected. Given adequate fit, the pattern of observed identification responses can reasonably be described in terms of the selected combination of d' and c values, enabling effects to be described in terms of differences in (estimates of) discriminability and bias.

In this study, the effects of feedback on discriminability and bias were assessed via analyses of d' and c estimates obtained using SDT-CD. Because SDT-CD is not in established use for the interpretation of identification data, the analyses of d' and c estimates were complemented by analyses of choosing rates and target-present accuracy rates for positive identifications.

Summary of the Present Research

A 2 (target presence) × 2 (feedback) + 2 (single-lineup: target presence) design was used to investigate whether feedback following an initial identification test influenced identification performance on a second test. Figure 1 shows an outline of the procedure. All participants viewed a simulated crime and, later, attempted to identify the culprit. In the multiple-lineup conditions, participants viewed two lineups: an initial target-absent lineup, and a second lineup that was target-present or target-absent and did not contain any foils from the initial target-absent lineup. (Target presence for the initial lineup was not manipulated as a condition in which participants viewed two target-present lineups would violate our requirement that no lineup members were to appear in both lineups, as outlined previously.) In the multiple-lineup conditions, feedback about the accuracy of the initial identification decision was manipulated prior to witnesses viewing the second lineup. Feedback comprised a statement that the initial lineup had not contained the culprit and the correct response to that lineup was "not present."

A single-lineup control condition was included to examine whether any effects of feedback following an initial identification test resulted in identification performance that differed from that of witnesses who viewed only one lineup. Participants allocated to the single-lineup control condition viewed either a target-present

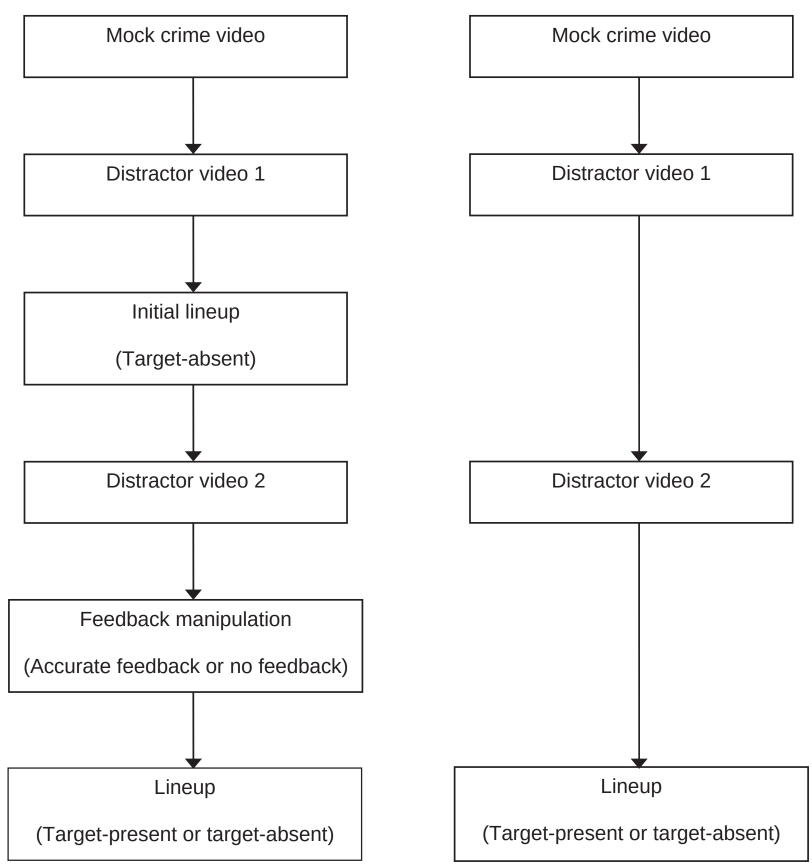

Figure 1. An outline of the procedure for participants in the multiple-lineups (left panel) and single-lineup (right panel) conditions.

or target-absent lineup. These lineups were the same as the target-present and target-absent lineups used for the second lineup in the multiple-lineups conditions. Participants in the single-lineup condition did not view any members of the initial, target-absent lineup used in the multiple-lineup conditions.

Method

Participants

Participants comprised 621 (387 female; aged 17 to 60 years, $M = 22.05$ years, $SD = 5.95$ years) undergraduate students and others recruited from interested community groups. All participants were paid an honorarium for their time.

Lineups

Head-and-neck photographs of the target and 12 match-description foils were used to construct three 6-person lineups. A target-present lineup (comprising the target and five foils) and target-absent lineup (including the same five foils, with the target replaced by another foil) were used for the single-lineup control condition and as the second lineup for the multiple-lineup conditions. Another target-absent lineup (comprising a different set of six foils) served as the initial lineup in the multiple-lineup conditions.

Following Palmer, Brewer, McKinnon, and Weber (2010), the suitability of foils was assessed as follows. First, one group of mock witnesses ($N = 5$) viewed the stimulus video and provided a description of the culprit. These descriptions were used to produce a modal description. Then, a second group ($N = 22$), that had not viewed the stimulus video, was presented with the modal description and photos of each of the 12 foils. This group was asked to indicate any faces that matched the description. On average, participants endorsed 10.27 ($SD = 1.72$) foils as matching the description. Each of the 12 foils was selected by between 68% and 100% of participants ($M = 85.61\%$, $SD = 11.27\%$).

Procedure

Upon arrival, participants were randomly allocated to an experimental or control condition and seated in an individual cubicle at a computer, via which all instructions were administered. Progress throughout the study was self-paced, with participants using a mouse to enter responses and navigate through the experiment. All participants viewed the stimulus video (which depicted a girl drinking coffee before stealing a male customer's wallet from a table at a café) and then a 5-min distractor video. After viewing the videos, participants in the multiple-lineup conditions were informed that they would be shown a lineup, and were asked to look for the girl who stole a wallet from the café in the first film. Participants were informed that the lineup may or may not contain this person, and were asked to respond by clicking on the appropriate photograph if the person was in the lineup, or clicking the "not present" button if the person was not in the lineup. Participants were then shown a target-absent lineup. Response latency was recorded via computer and participants were asked to rate their confidence that they had made a correct decision on an 11-point scale (ranging from 0%, *not at all confident,* to 100%, *completely confident*). Participants then viewed a second distractor video, after which they were informed that they would be shown another lineup that, again, may or may not contain the girl who stole a wallet from a cafe in the first video. Participants were told that the lineup they were about to see would not contain any person who had appeared in the first lineup, and were again reminded that the person they were looking for may or may not be in the lineup. Participants randomly allocated to the feedback condition were also informed that the lineup they saw previously did not contain the girl who stole a wallet at a café, and that the correct response to the initial lineup was "not present." Participants then viewed and responded to a target-present or target-absent lineup and rated their confidence in their decision as per the initial lineup.

Following the stimulus video and initial distractor video, participants in the single-lineup control conditions viewed the second distractor film, received lineup instructions (as per the initial lineup above), and viewed a target-present or target-absent lineup. Identification responses and confidence ratings were made as per the multiple lineup conditions. Note that participants in the single-lineup control conditions did not view any of the foils used for the initial target-absent lineup in the multiple-lineup conditions. In addition, because only one identification test was performed, the retention interval was approximately one minute shorter than the multiple-lineup conditions. There is no evidence from the eyewitness literature that this difference is important. Upon completion of the study, all participants were thanked and debriefed.

Results

The Effects of the Feedback Manipulation

Estimates of discriminability (d') and response bias (c). Table 1 provides an overview of the proportions and frequencies of identification responses made to target-present and target-absent lineups for each combination of initial response and feedback condition. Note that these data do not include responses to the initial target-absent lineup used in the multiple lineup conditions. An alpha level of .05 was used for all analyses. As cut offs for small, medium, and large effects, Cohen (1988) recommended using values of 0.1, 0.3, and 0.5, respectively, for the effect size estimate *w* (for comparisons of proportions), and 0.1, 0.25, and 0.4 for the effect size estimate *f* (for comparisons of means).

In order to identify and assess the best-fitting combination of d' and *c* values for each condition, observed and model-generated response probabilities (see Table 2) were compared using likelihood ratio *G*-tests (Sokal & Rohlf, 1981). These are conceptually very similar to χ^2 tests, but have an advantage in that the *G* statistic can be summed across tests. For each condition, three separate *G*-tests were conducted to compare observed and model-generated expected frequencies of false alarms, hits, and correct identifications. The total *G*-statistic for the three tests was calculated, with the smallest total *G* value indicating the best-fitting combination of d' and *c* estimates for each condition (see Table 3). For all conditions, the selected model adequately fit the data; all total $G\text{s}(df = 3) < 1.60$. Additional details of the statistical procedures used to calculate SDT-CD estimates of d' and *c* are included in the Appendix.

A bootstrap procedure was used to generate estimates of variance for d' and *c* values (e.g., Efron, 1981; Efron & Gong, 1983).

Table 1
Percentages and Frequencies of Identification Responses for Target-Present and Target-Absent Lineups

Initial decision and feedback condition	Identification response: Target-present lineups						
	Correct identification		Foil identification		Lineup rejection		Total
	%	*n*	%	*n*	%	*n*	*n*
Choosers							
No feedback	17.1	6	14.3	5	68.6	24	35
Feedback	10.3	3	44.8	13	44.8	13	29
Overall	14.1	9	28.1	18	57.8	37	64
Non-choosers							
No feedback	27.1	26	20.8	20	52.1	50	96
Feedback	40.8	42	6.8	7	52.4	54	103
Overall	34.2	68	13.6	27	52.3	104	199
Overall							
No feedback	24.2	32	19.1	25	56.5	74	131
Feedback	34.1	45	15.2	20	50.8	67	132
Single lineup	38.5	20	11.5	6	50.0	26	52
Total	30.8	97	16.2	51	53.0	167	315

Initial decision and feedback condition	Identification response: Target-absent lineups				
	Foil identification		Correct rejection		Total
	%	*n*	%	*n*	*n*
Choosers					
No feedback	28.2	11	71.8	28	39
Feedback	65.4	17	34.6	9	26
Overall	43.1	28	56.9	37	65
Non-choosers					
No feedback	31.5	28	68.5	61	89
Feedback	21.6	22	78.4	80	102
Overall	26.2	50	73.8	141	191
Overall					
No feedback	30.5	39	69.5	89	128
Feedback	30.5	39	69.5	89	128
Single lineup	18.0	9	82.0	41	50
Total	28.4	87	71.6	219	306

Following Weber and Brewer (2006), these estimates of variance were used to create inferential 95% confidence intervals (Tryon, 2001) for the d' and c values for each condition (shown in Table 3), allowing inferential comparisons to be made. Nonoverlapping confidence intervals indicate a statistically significant difference (at the $\alpha = .05$ level) between conditions. (See the Appendix for additional details of the procedures used for bootstrapping and the calculation of inferential 95% confidence intervals.)

Higher positive values of d' indicate better ability to distinguish the target from foils, with a value of zero indicating no ability to discriminate between targets and foils. For negative values of d', greater magnitude indicates increasing discriminability in the wrong direction (i.e., with foils selected more often than the target). A c value of zero indicates unbiased responding, with positive and negative values representing conservative and lenient biases (relative to unbiased responding), respectively. Consistent with our hypotheses, inspection of the inferential confidence intervals for d' values indicates that, relative to no feedback, feedback was associated with lower estimated d' for initial choosers and higher estimated d' for initial nonchoosers. Although there appeared to be a trend toward feedback being associated with more lenient responding among initial choosers, inspection of the inferential confidence intervals for c values suggests that response bias did not reliably vary with feedback for initial choosers or non-choosers.

Choosing and target-present accuracy rates. In the next section, we report that the conclusions drawn from analyses of SDT-CD estimates of d' and c align very closely with those obtained from analyses of more conventional measures of eyewitness identification performance (i.e., choosing and accuracy rates). It should be noted that drawing inferences about differences in discriminability and bias from choosing and accuracy data is often not straightforward (Duncan, 2006, 2010). In the present study, evidence for effects of feedback on discriminability would be provided by either of two patterns of results. First, if feedback (vs. no feedback) is associated with a difference in choosing rate from target-present or target-absent lineups, but not both, it would suggest that feedback influences detection discriminability (e.g., higher choosing from target-present lineups only would indicate greater detection discriminability). Second, if feedback (vs. no feedback) is associated with a difference in the ratio of correct identifications to foil identifications from target-present line-

Table 2
Observed and (Best-Fitting) Model-Generated Response Proportions for Target-Present and Target-Absent Lineups

Initial decision and feedback condition	Target-present: Correct identification	Target-present: Foil identification	Target-present: Lineup rejection	Target-absent: Foil identification	Target-absent: Lineup rejection
Choosers					
No feedback					
Observed	.19	.15	.67	.27	.73
Model	.19	.20	.62	.23	.77
Feedback					
Observed	.09	.48	.44	.70	.30
Model	.07	.54	.39	.65	.35
Non-choosers					
No feedback					
Observed	.25	.25	.51	.39	.61
Model	.24	.28	.48	.36	.64
Feedback					
Observed	.41	.09	.51	.23	.78
Model	.42	.10	.48	.20	.80
Single lineup					
Observed	.33	.13	.53	.18	.82
Model	.33	.13	.55	.19	.81

ups, it would suggest that feedback affects SDT-identification discriminability (e.g., a greater ratio of correct identifications to foil identifications would indicate greater discriminability). Evidence of feedback effects on response bias would be found if feedback (vs. no feedback) was associated with higher or lower choosing rates for both target-present and target-absent lineups, provided that the ratio of correct identifications to foil identifications from target-present lineups did not differ between feedback conditions.

A 2 (target-presence) × 2 (initial response) × 2 (feedback) × 2 (choosing status) hierarchical log-linear analysis was conducted, excluding data from the single-lineup conditions. This yielded a 3-way association between initial response, feedback, and choosing status, $\chi^2(1, n = 519) = 12.58$, $p < .01$. Follow-up chi-square analyses revealed that, for initial choosers, feedback was associated with a greater rate of positive identifications (62.8% with feedback vs. 29.8% without), $\chi^2(1, n = 129) = 11.82$, $p < .01$, $w = 0.30$. In contrast, for initial nonchoosers, the choosing rate did not differ significantly between those who received feedback (34.6%) and those who did not (40.0%), $\chi^2(1, n = 390) = 1.20$, *ns*, $w = 0.06$. Target-presence did not moderate the relationship between feedback and choosing rate for initial choosers, $\chi^2(1, n = 129) < 1$, or nonchoosers, $\chi^2(1, n = 390) = 1.31$, *ns*.

A 2 (initial response) × 2 (feedback) × 2 (accuracy) hierarchical log-linear analysis was conducted on positive identifications from target-present lineups, again excluding data from the single-lineup conditions. This yielded a significant 2-way initial response × accuracy association, $\chi^2(1, n = 122) = 14.01$, $p < .01$, $w = 0.33$, with fewer correct identifications made by initial choosers (33.3%) than initial nonchoosers (71.6%). Of greater interest was the significant 3-way association, $\chi^2(1, n = 122) = 11.08$, $p < .01$. Follow-up chi-square analyses showed that, for initial choosers, feedback was associated with a moderately smaller proportion of correct identifications (18.8% with feedback vs. 54.5% without), $\chi^2(1, n = 27) = 3.76$, $p = .05$, $w = 0.37$. In contrast, for initial nonchoosers, feedback was associated with a moderately

Table 3
Estimates of Discriminability (d′) and Response Bias (C), With Estimated Standard Errors in Parentheses, and Inferential 95% Confidence Intervals (CI)

Initial decision and feedback condition	Estimated discriminability: d'	95% CI Lower	95% CI Upper	Estimated response bias: c	95% CI Lower	95% CI Upper
Choosers						
No feedback	1.02 (0.37)	0.50	1.54	0.23 (0.20)	−0.05	0.51
Feedback	−0.08 (0.38)	−0.62	0.46	−0.23 (0.22)	−0.54	0.08
Non-choosers						
No feedback	1.29 (0.19)	1.03	1.55	−0.12 (0.11)	−0.27	0.03
Feedback	2.27 (0.21)	1.97	2.57	−0.24 (0.11)	−0.39	−0.09
Single lineup	2.09 (0.30)	1.66	2.52	−0.17 (0.16)	−0.39	0.06

larger proportion of correct identifications (85.7% with feedback vs. 56.5% without), $\chi^2(1, n = 95) = 9.94$, $p < .01$, $w = 0.32$.

For initial choosers, these results indicate that discriminability (in terms of the SDT-identification component of lineup decisions) was poorer with feedback than without feedback. It is important that although the choosing data suggest that feedback (vs. no feedback) was associated with more lenient responding, the target-present accuracy data rule out the possibility that feedback only influenced response bias. For initial nonchoosers, the results indicate that feedback (vs. no feedback) was associated with greater discriminability (in terms of SDT-identification) but had no significant effect on response bias.

Thus, regarding the effects of feedback on discriminability, the analyses of choosing and target-present accuracy rates converged with the analyses of d' estimates. Both sets of analyses indicate that feedback (vs. no feedback) was associated with poorer discriminability for initial choosers and better discriminability for initial nonchoosers. The effects of feedback on response bias were not as clear-cut. For initial nonchoosers, both choosing rates and c estimates suggest that feedback did not influence response bias. However, for initial choosers, the choosing data suggest that feedback was linked to a lenient shift in bias, but the analyses of c values indicated no reliable effect.

Response latency. A 2 (initial response) × 2 (feedback) ANOVA on logarithmic response latency data for the second lineup yielded no significant main effects or interaction, $Fs(1, 356) < 1.63$, $fs < 0.09$.

Comparisons With the Single-Lineup Control Condition

Estimates of discriminability (d') and response bias (c). To assess whether the effects of feedback following an initial identification test resulted in identification responses that differed from those of witnesses who viewed only one lineup, estimates of d' and c for the single-lineup control condition (see Table 3) were compared with those for each of the multiple-lineup conditions.

For estimated c values, the inferential confidence intervals for each of the multiple-lineup groups overlapped that for the single-lineup control condition, suggesting that response bias in each of the multiple-lineup conditions did not vary reliably from that in the control condition. However, the inferential confidence intervals indicate that estimated d' was greater for the single-lineup control condition than for initial choosers who received no feedback, initial choosers who received feedback, and initial nonchoosers who received no feedback. There was no significant difference in estimated d' between controls and initial nonchoosers who received feedback. These results suggest that, compared with witnesses who viewed only one lineup, discriminability was no different for initial nonchoosers who received feedback but was reliably inferior for all other multiple-lineup conditions.

Accuracy rates. From an applied perspective, we were interested in whether the differences in discriminability between the control condition and three of the multiple-lineup groups would translate to differences in overall identification accuracy. To investigate this, a series of 2 (condition) × 2 (target-presence) × 2 (accuracy) hierarchical log-linear analyses were conducted to compare the control condition with each of the multiple-lineup groups. Note that omnibus analyses were not performed here as the control group could not meaningfully be allocated a level on each of the independent variables (i.e., initial decision and feedback).

Compared with the single lineup condition (59.8%), the overall accuracy rate was lower for initial choosers who received no feedback (45.9%), $\chi^2(1, n = 176) = 5.55$, $p < .05$, $w = 0.14$; initial choosers who received feedback (21.8%), $\chi^2(1, n = 157) = 24.76$, $p < .01$, $w = 0.36$; and initial nonchoosers who received no feedback (47.0%), $\chi^2(1, n = 287) = 4.96$, $p < .05$, $w = 0.12$. The accuracy rate did not differ significantly between controls and initial nonchoosers who received feedback (59.5%), $\chi^2(1, n = 307) < 1$, *ns*, $w = 0.00$. None of the three-way associations were significant (all χ^2s < 1), indicating that differences in accuracy between the single lineup condition and other groups did not vary with target-presence. Thus, as per the analyses of d' estimates, the accuracy data indicate that identification performance of all multiple-lineup conditions (except initial nonchoosers who received feedback) was inferior to that of controls. These results will be interpreted in more detail in the Discussion.

Follow-Up Study

The results of this experiment suggest that, as predicted, feedback influenced discriminability on a second identification test. According to our proposed working account of the mechanism underpinning this relationship, these effects of feedback on discriminability operate via differences in witnesses' metacognitive beliefs (e.g., about their ability to perform an identification test involving the culprit). We conducted a simple follow-up study ($N = 133$, 57 female; aged 17 to 64 years, $M = 28.38$ years, $SD = 14.11$ years) to further test this idea. The procedure was very similar to that used for the previously reported experiment, but the primary dependent measure was witnesses' ratings of preidentification confidence (i.e., predicted likelihood that they would make a correct response on an upcoming second identification test involving the same culprit), rather than identification responses to a second test. Preidentification confidence ratings were used to index witnesses' metacognitive beliefs as there is evidence that performance predictions can be taken to reflect perceptions of ability to perform a specific upcoming memory task (e.g., Hertzog, Dixon, & Hultsch, 1990). If postidentification feedback affects discriminability via differences in witnesses' metacognitive beliefs, the effects of feedback on preidentification confidence should follow the same pattern as the effects of feedback on discriminability. That is, feedback (vs. no feedback) should be associated with lower preidentification confidence for initial choosers and higher preidentification confidence for initial nonchoosers. Ideally, the effects of feedback on preidentification confidence ratings and identification responses would be examined in a single experiment. However, we chose to conduct separate experiments because there is some evidence that measuring preidentification confidence substantially reduces identification accuracy (Fleet, Brigham, & Bothwell, 1987).

Results. Means and standard deviations for preidentification confidence ratings are shown in Table 4. A 2 (initial response) × 2 (feedback) ANOVA conducted on preidentification confidence ratings (excluding the control condition) yielded a significant and large main effect of initial decision, $F(1, 109) = 26.02$, $p < .01$, $f = 0.47$, with preidentification confidence greater among initial nonchoosers than initial choosers. There was no main effect of

Table 4
Mean Pre-Identification Confidence Ratings (With Standard Deviations in Parentheses) for the Follow-Up Study

	Feedback condition		
Initial decision	No feedback	Feedback	Overall
Choosers	56.67 (22.21)	44.00 (26.69)	51.21 (24.84)
Non-choosers	61.88 (16.41)	78.00 (13.63)	70.96 (16.82)
Overall	58.86 (19.98)	62.82 (26.52)	60.82 (23.43)
Single lineup			76.00 (11.07)

feedback, $F(1, 76) = 0.22$, *ns*, $f = 0.09$, but the interaction was significant, $F(1, 109) = 14.03$, $p < .01$, $f = 0.37$. Simple effects analyses indicated that, for initial choosers, preidentification confidence was lower following feedback compared with no feedback, $t(56) = 1.97$, $p = .05$, $f = 0.26$. In contrast, for initial nonchoosers, preidentification confidence was higher following feedback rather than no feedback, $t(53) = -3.98$, $p < .01$, $f = 0.54$. Thus, the effects of feedback on preidentification confidence followed the same pattern as the effects on discriminability (as evidenced by estimated d' values, choosing and accuracy rates) found in the previous experiment.

A series of independent samples *t* tests was conducted to determine whether preidentification confidence differed between witnesses yet to view a lineup and those who had performed an initial identification test. (Again, omnibus analyses were not performed as the control group could not be allocated a level on each of the independent variables.) In comparison with the control group, preidentification confidence was lower for initial choosers who received no feedback, $t(49.59) = 4.21$, $p < .01$, $f = 0.60$; initial choosers who received feedback, $t(33.48) = 5.44$, $p < .01$, $f = 0.85$; and initial nonchoosers who received no feedback, $t(42) = 3.28$, $p < .01$, $f = 0.52$. There was no significant difference in preidentification confidence between witnesses in the control condition and those who received feedback after an initial correct rejection, $t(49) = -0.55$, *ns*, $f = 0.08$. Again, this pattern of results mapped closely onto that obtained from identification responses in the previous experiment where, compared with the control condition, estimated discriminability and overall accuracy were on par for initial nonchoosers who received feedback but lower for all other groups.

Discussion

In this study, we tested whether accurate feedback following an initial identification decision would affect witnesses' responses to a second lineup for the same culprit. Contrary to what many recognition memory researchers might have expected, there was little evidence that feedback influenced response bias in a manner consistent with prior research (e.g., Stretch & Wixted, 1998, Experiment 1). In particular, there was no support for the idea that confirming feedback would produce a conservative shift in bias for initial nonchoosers. Instead, the results indicated that feedback influenced discriminability on a second test. For witnesses who chose a foil from an initial target-absent lineup, disconfirming feedback was associated with poorer discriminability on a second identification test (as evidenced by a smaller estimated d' and smaller ratio of correct identifications to foil identifications from target-present lineups, compared with the no-feedback condition). In contrast, for witnesses who correctly rejected the initial lineup, confirming feedback was associated with greater discriminability (as evidenced by a larger estimated d' and larger ratio of correct identifications to foil identifications from target-present lineups, compared with the no-feedback condition).

These results have implications not only for the eyewitness identification literature, but also for the broader recognition memory literature. The presence of feedback effects on discriminability highlights the importance of acknowledging that decision making in recognition memory involves more than simply setting a criterion for the amount of evidence required for a positive response. This study adds to a growing body of theory and empirical work demonstrating that people are not inflexible in terms of the way that they use different types of evidence when making recognition decisions (e.g., Banks, 2000; Herron & Rugg, 2003; Jacoby et al., 2005; Lane, Roussel, Starns, Villa, & Alonzo, 2008; Lane et al., 2007; Starns et al., 2007; Whittlesea & Williams, 1998).

In terms of the mechanism underpinning the effects of feedback on discriminability, we have outlined one account that draws on the source monitoring framework (Johnson et al., 1993), multidimensional models of memory (e.g., Banks, 2000), and related theory and empirical work on the distinctiveness heuristic (e.g., Dodson & Schacter, 2001, 2002; Schacter et al., 1999). According to this account, postidentification feedback influences witnesses' metamnemonic beliefs (e.g., regarding their ability to perform identification tests involving the culprit). The resulting differences in metamnemonic beliefs affect the way that witnesses evaluate memorial evidence during a second identification decision, such that confirming feedback is associated with more stringent evaluation and disconfirming feedback with less stringent evaluation. As a result, confirming feedback enables initial nonchoosers to better distinguish between the culprit and foils in a second lineup. Conversely, disconfirming feedback impairs initial choosers' ability to distinguish the culprit from foils in a second lineup.

This account received some support from the results of the follow up-study, which showed that disconfirming feedback was associated with lower preidentification confidence ratings (for initial choosers) and confirming feedback was associated with higher preidentification confidence ratings (for initial nonchoosers). However, several elements of the proposed mechanism remain untested. For example, although these results suggest that the effects of feedback on witnesses' metacognitive beliefs were as predicted, we cannot be certain that these differences in metacognitions caused witnesses to alter the way that memory information was used during a second identification decision. Moreover, as per Lane et al. (2007), we make no claims about how witnesses might have used memorial information differently during the second identification decisions. For example, feedback may have prompted witnesses to weight various types of information differently, ignore some types entirely, or take new types of information into account. Thus, we emphasize that this account is proposed as a useful working hypothesis only.

As an aside, it is important to note that the different effects of feedback for initial choosers and nonchoosers cannot be explained by differences in memory quality. Because responses to the initial target-absent lineup were correct for initial nonchoosers and incorrect for initial choosers, it could be argued that initial nonchoos-

ers had, on average, better memories for the target than initial nonchoosers. This likely difference in memory quality might explain the overall difference in performance on the second identification test (e.g., higher target-present accuracy for initial nonchoosers than choosers). However, the effects of feedback were examined by comparing the feedback and no feedback conditions within each type of initial response (i.e., initial choosers or nonchoosers), and there is no reason to believe that memory quality differed between initial choosers (or nonchoosers) who received feedback and those who did not.

In developing an understanding of the effects of feedback on subsequent identification performance, we also have to take into account the results of comparisons between the multiple-lineup and single-lineup conditions. Compared to witnesses who viewed only one lineup, estimated discriminability and overall accuracy were equivalent for initial nonchoosers who received feedback but poorer for all other multiple lineup groups. Overall, these results suggest that the use of an initial identification test has the potential to impair, but not enhance, performance on a second identification test. When considering the multiple lineup conditions separately, the inferior performance of initial choosers compared with controls was not surprising, as it has been suggested that choosing a foil from an initial lineup likely indicates relatively poor memory for the culprit (Wells, 1984). However, we did not anticipate that initial nonchoosers who received no feedback would perform worse than controls. This result implies that correctly rejecting the initial lineup in the absence of feedback actually impaired performance on the second identification test. The results of the follow-up study suggest that this difference in performance was accompanied by a difference in metacognitive beliefs. Compared with the single-lineup control condition, preidentification confidence was lower for initial nonchoosers who received no feedback ($f = 0.52$). Thus, rejecting the initial lineup (in the absence of feedback) apparently had a negative influence on witnesses' perceptions of their ability to perform a second identification test for the same culprit.

There are at least two explanations for how this could occur. First, witnesses may have found the process of viewing and rejecting the initial lineup more difficult than they expected to. If so, the subjective difficulty experienced may well have negatively influenced their perceived ability to perform a second identification test. Second, some witnesses may have approached the initial identification test expecting that the culprit would be present and that they would be able to identify that person (cf. Wells, 1993). Such witnesses may have viewed their initial lineup rejection as a failure to adequately perform the identification task, rather than an appropriate response to a lineup that might not contain the culprit. Consequently, they would likely have felt less confident about responding correctly to a second lineup.

The results of the control comparisons also shape our interpretation of the beneficial effects of confirming feedback for initial nonchoosers. Discriminability, target-present accuracy, and (in the follow-up study) preidentification confidence were very similar across the control condition and initial choosers who received feedback. This suggests that accurate confirming feedback merely counteracted the negative influences on witnesses' metacognitive beliefs resulting from the initial lineup rejection. For example, after learning that their initial lineup rejection was correct, witnesses would no longer have reason to think that their initial decision represented a failure to adequately perform the identification test.

From an applied perspective, the results of the control group comparisons have clear implications for criminal investigative procedures. As noted earlier, these results suggest that the use of an initial identification test has the potential to impair, but not enhance, performance on a second identification test. Thus, for investigations of single-culprit crimes, the use of multiple lineups can potentially result in the conviction of more innocent suspects and fewer guilty suspects. The simplest strategy for avoiding the use of multiple lineups is for investigators to delay conducting an identification test until there is substantial evidence as to the guilt of one particular suspect. However, if this is not possible (e.g., in cases where there are multiple plausible suspects for a single-culprit crime), an alternative approach would be to place the most likely suspect (i.e., the one to whom the most additional evidence points) in the first lineup to be viewed. Although a suspect identification from the first lineup does not necessarily indicate guilt, our results suggest that it would be more reliable than one made from a subsequent lineup.

The findings regarding the effects of feedback also have clear applied value for cases where the use of multiple lineups cannot be avoided. First, confirming feedback following a correct lineup rejection improved identification performance. Here, it is important to note that our results do not in any way suggest that providing confirming feedback after a suspect identification will improve identification performance. The results do, however, suggest that if a witness rejects an initial lineup, investigators should refrain from showing a second lineup unless they can inform the witness that their initial lineup rejection was correct (e.g., via new evidence pointing to the innocence of the first suspect). Second, disconfirming feedback following an incorrect positive identification reduced performance. Thus, if a witness makes a positive identification from an initial lineup (i.e., identifies a foil or a suspect who is later proven innocent) investigators should avoid allowing the witness to learn that their initial choice was incorrect. This is potentially difficult because, even in the absence of explicit feedback, disconfirming feedback might be inferred by witnesses (e.g., "If they are asking me to come back to the police station to view another lineup, the person I picked last time must not have been the culprit"). One solution might be for investigators to ask the witness to view any subsequent lineups as a matter of standard procedure.

Finally, this study represents a test of the utility of SDT-CD (Duncan, 2006, 2010; see also Starr et al., 1975) for calculating estimates of discriminability and response bias for compound decisions. The conclusions drawn from analyses of SDT-CD-derived estimates of d' and c were almost identical to those drawn from analyses of choosing and accuracy rates. If results obtained with SDT-CD can be shown to reliably converge with those obtained via analyses of more established metrics, this model will likely prove very useful for the efficient interpretation of effects in studies involving compound decisions.

In sum, this research provides evidence that feedback following an initial identification test can affect performance on a second test involving the same culprit. From a theoretical viewpoint, the results are consistent with the idea that the way people use memorial information when making memory decisions is flexible. From an applied perspective, the findings imply that the use of

multiple lineups can impair, but not enhance, subsequent identification accuracy. Hence, the use of multiple lineups should be avoided in criminal investigations.

References

Anderson, H. E., White, W. F., & Wash, J. A. (1966). Generalized effects of praise and reproof. *Journal of Educational Psychology, 57,* 169–173.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11,* 267–273.

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

Bouffard-Bouchard, T. (1993). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology, 30,* 353–363.

Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology, 87,* 112–120.

Bridgeman, B. (1974). Effects of test score feedback on immediately subsequent test performance. *Journal of Educational Psychology, 66,* 62–66.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. New York: Cambridge University Press.

Davies, G., Shepherd, J., & Ellis, H. (1979). Effects of interpolated mugshot exposure on accuracy of eyewitness identification. *Journal of Applied Psychology, 64,* 232–237.

Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30,* 287–307.

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin and Review, 8,* 155–161.

Dodson, C. S., & Schacter, D. L. (2002). When false recognition meets metacognition: The distinctiveness heuristic. *Journal of Memory and Language, 46,* 782–803.

Duncan, M. J. (2006). *A signal detection model of compound decision tasks* (Tech. Rep. No. No. TR2006–256). Toronto, ON: Defence Research and Development Canada.

Duncan, M. J. (2010). *A signal detection model of compound decision tasks.* Manuscript submitted for publication.

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap, and other resampling methods. *Biometrika, 68,* 589–599.

Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician, 37,* 36–48.

Feather, N. T. (1966). Effects of prior success and failure on expectations of success and subsequent performance. *Journal of Personality and Social Psychology, 3,* 287–298.

Feather, N. T., & Saville, M. R. (1967). Effects of amount of prior success and failure on expectations of success and subsequent task performance. *Journal of Personality and Social Psychology, 5,* 226–232.

Fleet, M. L., Brigham, J. C., & Bothwell, R. K. (1987). The confidence-accuracy relationship: The effects of confidence assessment and choosing. *Journal of Applied Social Psychology, 17,* 171–187.

Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recollection tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory & Language, 51,* 473–493.

Godfrey, R. D., & Clark, S. E. (2010). Repeated eyewitness identification procedures: Memory, decision making, and probative value. *Law and Human Behavior, 34,* 241–258.

Gronlund, S. D. (2004). Sequential lineups: Shift in criterion or decision strategy? *Journal of Applied Psychology, 89,* 362–368.

Halford, P. (2009). *Eyewitness identification in real world conditions.* Unpublished doctoral dissertation, Flinders University, Adelaide, South Australia, Australia.

Herron, J. E., & Rugg, M. D. (2003). Strategic influences on recollection in the exclusion task: Electrophysiological evidence. *Psychonomic Bulletin & Review, 10,* 703–710.

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5,* 215–227.

Hinz, T., & Pezdek, K. (2001). The effect of exposure to multiple lineups on face identification accuracy. *Law and Human Behavior, 25,* 185–198.

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 302–313.

Innocence Project. (2009). *Eyewitness misidentification.* Retrieved from http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review, 12,* 852–857.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Klein, S. B., Loftus, J., & Fricker, S. S. (1994). The effects of self-beliefs on repeated efforts to remember. *Social Cognition, 12,* 249–261.

Lane, S. M., Roussel, C. C., Starns, J. J., Villa, D., & Alonzo, J. D. (2008). Providing information about diagnostic features at retrieval reduces false recognition. *Memory, 16,* 836–851.

Lane, S. M., Roussel, C. C., Villa, D., & Morita, S. K. (2007). Features and feedback: Enhancing metamnemonic knowledge at retrieval reduces source-monitoring errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 1131–1142.

Lloyd, M. E., Westerman, D. L., & Miller, J. K. (2003). The fluency heuristic in recognition memory: The effect of repetition. *Journal of Memory & Language, 48,* 603–614.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

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,* 783–792.

Nettles, W., Nettles, Z., & Wells, G. L. (1996). "I noticed you paused on number three": Biased testing in eyewitness identification. *Champion, 10, 12,* 57–59.

Osler, S. F. (1954). Intellectual performance as a function of two types of psychological stress. *Journal of Experimental Psychology, 47,* 115–121.

Palmer, M. A., Brewer, N., McKinnon, A. C., & Weber, N. (2010). Phenomenological reports diagnose accuracy of eyewitness identification decisions. *Acta Psychologica, 133,* 137–145.

Pezdek, K., & Blandon-Gitlin, I. (2005). When is an intervening line-up most likely to affect eyewitness identification accuracy? *Legal and Criminological Psychology, 10,* 247–263.

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review, 12,* 865–873.

Schacter, D. L., Israel, L., & Racine, C. A. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language, 40,* 1–24.

Semmler, C., & Brewer, N. (2006). Postidentification feedback effects on face recognition confidence: Evidence for metacognitive influences. *Applied Cognitive Psychology, 20,* 895–916.

Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of postidentification feedback on eyewitness identification and nonidentification. *Journal of Applied Psychology, 89,* 334–346.

Smith, S. M., Lindsay, R. C. L., & Pryke, S. (2000). Postdictors of eyewitness errors: Can false identifications be diagnosed? *Journal of Applied Psychology, 85,* 542–550.

Sokal, R. R., & Rohlf, F. J. (1981). *Biometry* (2nd ed.). San Francisco: Freeman.

Starns, J. J., Lane, S. M., Alonzo, J. D., & Roussel, C. C. (2007). Metamnemonic control over the discriminability of memory evidence: A signal detection analysis of warning effects in the associative list paradigm. *Journal of Memory and Language, 56,* 592–607.

Starr, S. J., Metz, C. E., Lusted, L. B., & Goodenough, D. J. (1975). Visual detection and localization of radiographic images. *Radiology, 116,* 533–538.

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

Strack, F., & Forster, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science, 6,* 352–358.

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, and Cognition, 24,* 1379–1396.

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods, 6,* 371–386.

Weber, N., & Brewer, N. (2006). Positive versus negative face recognition decisions: Confidence, accuracy, and response latency. *Applied Cognitive Psychology, 20,* 17–31.

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

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

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83,* 360–376.

Wells, G. L., & Bradfield, A. L. (1999). Distortions in eyewitnesses' recollections: Can the postidentification feedback effect be moderated? *Psychological Science, 10,* 138–144.

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica, 98,* 141–165.

Appendix

Additional Details of Statistical Procedures

Calculation of estimated d′ and c. For the likelihood ratio *G*-tests (Sokal & Rohlf, 1981), expected frequencies were calculated by multiplying model-generated response probabilities by the number of target-absent and target-present trials for false alarms and hits, respectively, and by the number of observed hits in the case of correct identifications. A broad range of possible parameter values was considered (i.e., −1.59 to 4.01 for d' and −3 to 3 for decision criterion) in the model-selection process.

It should be noted that there are two versions of SDT-CD, based on different decision rules that decision-makers might adopt (Duncan, 2006, 2010). The *independent observation* rule assumes that each stimulus in the array is assessed separately against the decision criterion and a positive response is made if the strength of evidence for one or more stimuli exceeds the criterion. The *integration* rule assumes that decision-makers make a more global assessment of the array, comparing the sum of strength-of-evidence values for all stimuli against the decision criterion. Here, only estimates based on the integration decision rule are reported, as these provided a better fit than estimates based on the independent observation rule for every condition. While it is beyond the scope of this paper to address why this might be the case, it is worth noting that this pattern is consistent with previous analyses of eyewitness identification decisions (Duncan, 2006, 2010).

Calculation of inferential 95% confidence intervals. A bootstrapping procedure (e.g., Efron, 1981; Efron & Gong, 1983) was conducted on each condition separately. The response frequencies for each condition were used as a sampling distribution from which 250 replication data sets were randomly drawn. Estimates of d' and c values were calculated for each of these data sets, providing distributions of d' and c values from which estimates of variance could be derived for each parameter.

In the calculation of inferential 95% confidence intervals, E is a parameter that reflects the equivalence of standard errors across groups and influences the t value used to create inferential confidence intervals. We used an average E value, calculated across all pairwise combinations of standard errors included in the planned comparisons (Tryon, 2001).

Received January 28, 2010
Revision received June 25, 2010
Accepted July 13, 2010 ■