

The influence of retrieval processes in verbal overshadowing

CHRISTIAN A. MEISSNER, JOHN C. BRIGHAM, and COLLEEN M. KELLEY
Florida State University, Tallahassee, Florida

Recent studies of eyewitness memory have observed deleterious effects of producing a verbal description on later identification accuracy of a previously viewed face, an effect termed *verbal overshadowing* (Schooler & Engstler-Schooler, 1990). The present research investigated whether the phenomenon of verbal overshadowing may be constrained by variation in participants' initial retrieval processes, such that verbalization of a previously viewed stimulus could produce either positive or negative influences on subsequent attempts at recollection. To assess the validity of this hypothesis, we manipulated participants' response criterion during the verbal description task. As predicted, variation in response criterion significantly influenced not only the quality of the description generated but also accuracy on a subsequent identification task. This retrieval-based effect was found to persist despite either a postdescription delay (Experiment 1) or source-monitoring instructions at the time of the identification task (Experiment 2). We conclude that retrieval-based processes exert a powerful influence over the accuracy of verbalization and subsequent identification of a target face.

Standard law enforcement protocol has been to question eyewitnesses to a crime, often repeatedly over the course of an investigation, regarding various aspects of the scenario observed. In particular, obtaining a thorough description of the perpetrator of the crime is often crucial to the investigation, since it assists officers in quickly identifying and apprehending the suspect. Although person descriptions have been considered rather standard procedure, results of recent research have led the scientific community to question the influence of such a technique on the eyewitness's subsequent ability to provide an accurate identification.

The Verbal Overshadowing Effect

Schooler and Engstler-Schooler (1990) demonstrated that verbal description can have deleterious effects on later identification of a target face, an effect termed *verbal overshadowing*. After viewing a short video presentation of a bank robbery, participants either described the suspect from the scenario (description condition) or performed an innocuous task (no-description control condition). Recognition performance was then assessed using an 8-photo target-present lineup. Participants asked to provide a description of the target prior to identification were significantly less accurate than were controls on the identification task (38% and 64% accuracy, respectively, in Experiment 1).

Since their initial set of experiments, Schooler and Engstler-Schooler's (1990) general findings have been replicated within the facial memory paradigm (Dodson, Johnson, & Schooler, 1997; Fallshore & Schooler, 1995; Read & Schooler, 1994; Ryan, 1992; Ryan & Schooler, 1998; Schooler, Ryan, & Reder, 1996) and have been extended to other domains involving perceptual expertise (e.g., musical memory, Houser, Fiore, & Schooler, 1998; wine tasting, Melcher & Schooler, 1996). Schooler, Fiore, and Brandimonte (1997) reviewed these studies and further theorized on the processes involved in the verbal overshadowing effect. Specifically, the authors investigated three main factors believed to govern the effect: *modality mismatch* (i.e., the notion of competing verbal vs. nonverbal representations), *availability* (i.e., the assumption that the visual representation remains available/accessible in memory despite the temporary verbal interference), and *recoding interference* (i.e., the notion that overshadowing effects are due to a nonveridical verbal code of the target stimulus). Schooler et al. (1997) concluded that the majority of research appeared to provide strong support for both the modality mismatch and the availability assumptions. However, the recoding interference hypothesis lacked empirical support, primarily because studies within the facial memory domain have failed to find a relationship between description accuracy and identification accuracy (cf. Fallshore & Schooler, 1995; Schooler & Engstler-Schooler, 1990).

Retrieval-Based Effects

In contrast to the above pattern of findings, several attempts to replicate the verbal overshadowing effect in other labs have proven unsuccessful (Lindsay, personal communication, 1990, cited in Schooler et al., 1996; Lovett,

The authors thank Neil Charness, Anders Ericsson, Roddy Roediger, and Jonathan Schooler for their comments on earlier versions of this manuscript. Correspondence should be addressed to C. A. Meissner or J. C. Brigham, Department of Psychology, Florida State University, Tallahassee, FL 32306-1270 (e-mail: brigham@psy.fsu.edu).

Small, & Engstrom, 1992; Yu & Geiselman, 1993). For example, Lovett et al. attempted to directly replicate the findings from Experiment 1 of Schooler and Engstler-Schooler (1990). The description condition did not significantly differ from the control condition in all three of their experiments. Other studies have demonstrated positive effects of verbal description, rehearsal, and elaboration on later recognition of faces (Chance & Goldstein, 1976; Mauldin & Laughery, 1981; McKelvie, 1976; Read, 1979; Wogalter, 1991, 1996).

Schooler et al. (1996) commented on this seeming inconsistency in the literature by suggesting that failures to replicate may have resulted from variations in either (1) the encoding and retrieval processes in which participants spontaneously engaged or (2) the degree of interference generated by the description task. Whereas Schooler et al.'s comments were primarily focused on the notion of a process/modality mismatch (i.e., holistic/visual vs. featural/verbal processing of faces; see General Discussion), we propose that a return to Schooler and Engstler-Schooler's (1990) original *recoding interference* account, including consideration of more recent research on the influence of *retrieval-based effects*, may provide an alternative explanation. More specifically, we suggest that discrepancies across studies within the verbal overshadowing paradigm may be due to variation in the criterion that participants use as they attempt to retrieve their memory for the target during the description task.

In their review, Roediger and Gynn (1996) noted that variation in individuals' initial retrieval processes could significantly influence subsequent attempts at recollection. Such effects could be either positive (aiding later attempts at recollection) or negative (causing forgetting, interference, or false recollections), with the direction of influence largely depending on the accuracy of the initial retrieval instance. Early studies consistently demonstrated the biasing effects of *verbal labeling* (Carmichael, Hogan, & Walter, 1932; Herman, Lawless, & Marshall, 1957; also see Riley, 1962), whereas other more recent studies have demonstrated how *verbal misinformation* (Lindsay, 1990; Lindsay & Johnson, 1989; Loftus, 1975) can bias individuals' memory for faces (see Loftus & Ketcham, 1983). Although such misinformation is provided to individuals, several studies (Brigham & Cairns, 1988; Brown, Deffenbacher, & Sturgill, 1977; Gorenstein & Ellsworth, 1981) have shown that when participants themselves initially select an incorrect foil, they are more likely to incorrectly identify the same foil on a later recognition test, despite the addition of the original target face. Similarly, several studies have illustrated that initial recall/recognition of an incorrect item from a visual event significantly impairs memory on a subsequent test (Roediger, Jacoby, & McDermott, 1996; Schooler, Foster, & Loftus, 1988).

Roediger, Wheeler, and Rajaram (1993) also demonstrated the negative influences of an initial retrieval instance across repeated attempts at recall. They examined

whether participants requested to lower their response criterion ("forced" recall) would later be more susceptible to the effects of their own self-generated misinformation on a second memory test, relative to participants permitted to establish their own response criterion ("free" recall). Their results indicated that the forced recall condition was significantly more likely to generate unstudied items on both an immediate test and a delayed test when compared with a free recall condition. Roediger et al. concluded that forcing participants to generate misinformation by lowering their initial response criterion constituted a potent source of interference in later recall.

Ackil and Zaragoza (1998) similarly examined forced and free responding to both truthful and leading questions regarding a previously viewed event. When participants (first graders, third/fourth graders, and college students) returned 1 week later, they were instructed to monitor the source of their memories for both new and old questions about the event (including both truthful and leading phrasing). Ackil and Zaragoza found that, despite age differences in the magnitude of effects, forcing participants to confabulate responses on an initial test significantly increased the likelihood of confabulation and misattribution 1 week later. Furthermore, participants in the forced condition were more likely to later misattribute self-generated responses when compared with participants in the free recall condition who responded to experimenter-provided items.

Evidence of such *retrieval-based processes* within the verbal overshadowing paradigm has not been directly tested. However, in a related vein, Finger and Pezdek (1999) noted that variation in the type of description task significantly influenced later identification performance. Specifically, the authors were interested in whether the cognitive interview (CI; Fisher & Geiselman, 1992), an interview procedure found to elicit more verbal descriptions and critical information about a crime incident, would produce a verbal overshadowing effect when compared with a standard interview (SI) procedure. Previous studies investigating the influence of the CI on subsequent lineup identification had demonstrated no detrimental effects (Fisher, Quigley, Brock, Chin, & Cutler, 1990; Gwyer & Clifford, 1997); however, the degree to which the interview procedure specifically emphasized a facial description of the target was unknown. When the description of the perpetrator was made central to the interview procedure, Finger and Pezdek found that the CI did markedly impair recognition memory below that of the SI procedure. In addition, several follow-up experiments demonstrated that insertion of a postdescription delay significantly diminished or "released" the interference caused by the verbal overshadowing effect.

Finger and Pezdek (1999) also examined the role of participants' description accuracy on later identification and found that the relative precision of the verbal description was significantly influenced by the manner in which participants were instructed to recollect the infor-

mation during the description task (i.e., CI vs. SI). Furthermore, participants who misidentified the target had produced significantly more correct and incorrect details. These results appear consistent with the influence of retrieval-based processes and suggest that the representation created when participants generated a description of the target did not resemble the perceptual representation in memory. Finger and Pezdek proposed that the incongruent representation created during the description task may have remained more accessible initially, thereby influencing participants' immediate identification of the target.

EXPERIMENT 1

Taken together, the inconsistencies observed in the verbal overshadowing paradigm may stem from variations in the implicit demand for participants to adopt a lower or higher response criterion during the recall task. More specifically, it can be hypothesized that a lower response criterion may lead participants to generate more inaccuracies in their description, thereby subsequently producing the overshadowing effect on the identification task. In Experiment 1, we directly tested this hypothesis by inducing participants to generate descriptions using different response criteria. Additionally, we investigated the effect of varying the delay between description and identification phases on the "release" of verbal overshadowing, an issue previously addressed by Finger and Pezdek (1999).

The standard verbal overshadowing paradigm was followed. All participants initially viewed a target face, after which they were assigned to one of four conditions: (1) a *forced recall condition* in which participants' response criterion was lowered such that errors were likely to be generated when describing the target, (2) a *standard recall condition* in which participants were presented instructions patterned after those of previous verbal overshadowing experiments and were permitted to establish their own response criterion, (3) a *warning recall condition* in which participants' response criterion was raised such that errors were unlikely to be generated when describing the target, or (4) a *control condition* in which participants were not requested to provide a description of the target. Either immediately or following a 30-min delay, participants were presented with a target-present photo lineup from which they were asked to select the target.

We hypothesized that lowering participants' response criterion (forced recall condition) would increase the degree of erroneous information generated by participants, which, in turn, would reduce the likelihood of subsequent identification. Additionally, if the misinformation generated by participants acts as a source of retroactive interference in which access to the original perceptual memory is temporarily blocked, then insertion of a delay should eliminate this influence and subsequently increase identification accuracy (cf. Wheeler, 1995). Con-

versely, raising participants' response criterion (warning recall condition) was hypothesized to reduce the amount of erroneous information generated and thereby improve later recognition performance whether assessment was immediate or delayed. Finally, in attempting to replicate previous verbal overshadowing experiments, participants who were permitted to establish their own response criterion (standard recall condition) were hypothesized to demonstrate less accuracy on an immediate recognition task when compared with control participants who did not describe the target. However, when a postverbalization delay was inserted, participants who generated a description were expected to perform no differently than controls due to the "release" of retrieval-based interference that was created during the description task.

Method

Participants. The participants were 240 introductory psychology students (91 males and 149 females) who received course credit for their participation. They were tested in groups ranging in size from 2 to 8. Additionally, 28 students (9 males and 19 females) in an experimental psychology lab provided data for lineup fairness analyses.¹

Materials. The target photo and six-person photo lineup used by Finger and Pezdek (1999) were used here as stimulus materials. The target was a 31-year-old white male who faced a 45° angle. The six-person photo lineup consisted of full-frontal photos of the target individual and five foils of similar appearance. Two versions of the photo lineup were generated such that the target was presented in Position 3 and Position 4 an equal number of times.²

Design and Procedure. A 4 × 2 between-subjects design was used to examine the effects of variations in recall criterion (4) and postverbalization delay (2) on subsequent identification performance. Thirty participants were represented in each cell of the design. All participants initially viewed the target face for 10 sec. They were instructed that they would be shown a slide for several seconds and that, later, they would be asked questions about what they had viewed. Following the target presentation, all participants performed a 5-min filler task (digit search puzzle), after which they were assigned to one of three conditions in which manipulations were made to their response criterion as they were asked to describe the target. A fourth control group of participants did not participate in the description task but instead were asked to list as many of the American states as they could recall. The participants in all conditions were given 5 min to perform their task (i.e., description generation or listing of states).

Those who participated in the description task were given a response sheet with lines numbered from 1 to 25. All instructions were printed at the top of the response sheet and were read to participants prior to the description task. The participants in the standard recall condition were presented the following instructions:

In the spaces below, please describe in as much detail as possible the face you saw in the slide. Use the lines below to provide details about what the face looked like. You should attempt to describe the person in sufficient detail such that someone else could identify him on the basis of the description. As describing a face is often a difficult task, it is important that you concentrate and stay focused for the next few minutes.

In addition to the standard instructions, the participants in the warning recall condition were also told:

Prior research has also demonstrated the importance of striving for accuracy and reporting only that which you are certain you remember. You do not have to fill in all of the lines, so be sure to report only those details that you are confident of, and do not attempt to guess at any particular feature.

Finally, the participants in the forced recall condition were given the following instructions in addition to the standard instructions:

Prior research has also demonstrated the importance of reporting everything that you can remember about the individual on the slide. Try not to leave out any details about the face even if you think they are not important. You must fill in all of the numbered lines below with a description of the face, even if you start to feel that you are guessing.

Immediately following the description task, the participants in the 30-min delay condition were asked to assist the experimenter in another project. The participants were told that an attitude questionnaire was being developed concerning the practice of euthanasia and that they were to spend the next 30 min responding to the questionnaire. The participants were unaware that an identification task would later follow.

Either immediately or following the 30-min delay, the participants were presented with a photo lineup identification task. The participants were instructed that they would be shown a slide containing six faces and that they should choose the face that they saw at the beginning of the experiment. In addition, the participants were warned that the target may or may not be present in the slide and thus were given the option of not selecting a face from the photo lineup.

Results and Discussion

Identification accuracy. A 4 × 2 × 3 hierarchical loglinear analysis (HILOG) was performed to examine the influence of recall criterion (standard, forced, warning, or control condition) and postexposure delay (immediate or 30-min delay) on participants' identification accuracy (hit, miss, or false-alarm response). The HILOG analysis is analogous in design to that of the factorial analysis of variance (ANOVA); however, the HILOG analysis is more appropriate when both the independent and the dependent variables are categorical in nature. Results indicated a significant recall criterion × accuracy interaction [$\chi^2(6) = 21.88, p < .001$], such that systematic differences in identification accuracy were present among the description conditions. Neither the delay × accuracy [$\chi^2(2) = 0.48, n.s.$] nor the three-way recall criterion × delay × accuracy interaction [$\chi^2(2) = 1.62, n.s.$] were significant. Table 1 provides the proportion of identification responses across the levels of recall criterion and postdescription delay.

It is important to note that the description conditions were similar on the proportion of misses ("not present" responses) [$\chi^2(3) = 0.46, n.s.$] ($M = .17, SE = .02$), in-

dicating that differences between conditions were due to variations in the proportion of hits and false alarms generated by participants. Thus, performance on the identification task was a function of better or worse memory due to manipulations in the recall task, and not a bias against responding to the photo lineup.

To examine several a priori predictions, follow-up chi-square analyses were conducted. As predicted, participants in the forced condition performed significantly worse on the identification task than did participants in all other conditions [$\chi^2s(2) > 6.50, ps < .05$]. Additionally, participants in the warning condition significantly outperformed those in the forced and no-description control conditions [$\chi^2s(2) > 6.00, ps < .05$]. Finally, the typical verbal overshadowing effect was not replicated in the standard condition, since it did not significantly differ from the control condition [$\chi^2(2) = 1.41, n.s.$].

Description quality. To determine whether participants' description quality influenced the accuracy of identification responses observed across the recall criterion conditions, two independent coders examined each description for the number of correct, incorrect, and subjective details (cf. Finger & Pezdek, 1999). *Correct details* were defined as those facial aspects (eye color, hair texture, etc.) of the description that correctly matched the target. *Incorrect details* did not correctly match the target. *Subjective details* were those ambiguous qualities of face shape (e.g., round head, long face, etc.) or personality/occupational impressions (e.g., pleasant looking, a construction worker, etc.) that were not specific features. Coders trained to a criterion of 85% agreement before independently scoring each of the 180 descriptions. Interrater reliabilities were sufficient (correct details, $r = .92$; incorrect details, $r = .94$; subjective details, $r = .89$). For each participant, an estimate was obtained by averaging the coders' ratings within each category.

We posited that variation in the type of recall criterion presented to participants would significantly influence subsequent description accuracy. To evaluate this hypothesis, a multivariate analysis of variance (MANOVA) examined variations in the type of recall criterion (standard, forced, or warning recall conditions) on participants' overall description quality (number of correct, incorrect, and subjective details generated). The initial multivariate test indicated a significant main effect of type of recall criterion [$F(6,350) = 32.23, p < .001, \eta^2 = .36$]. Follow-up univariate tests indicated that the description conditions significantly differed in the number of correct, incorrect, and subjective details [$Fs(2,177) > 8.00, MS_{\epsilon}s = 4.36, 1.04, \text{ and } 3.68, \text{ respectively, } ps < .001$], confirming that manipulation of participants' response criterion did significantly influence overall description quality. Simple-effects comparisons (Bonferroni correction) indicated that all three description conditions significantly differed from one another in the number of both correct and subjective details [$ts(118) > 1.87, ps < .001$]. However, only the warning condition differed from both the forced condition and the standard condition in the number of incorrect details [$ts(118) > 3.60, ps < .001$].

Table 1
Proportion Accuracy as a Function of Type of Recall Criterion and Postdescription Delay in Experiment 1

	Control	Standard	Warning	Forced
	Immediate			
Hit	.47	.47	.67	.27
Miss	.17	.17	.13	.13
False alarm	.36	.36	.20	.60
	30-min Delay			
Hit	.37	.57	.60	.27
Miss	.30	.17	.10	.17
False alarm	.33	.26	.30	.56
	Overall			
Hit	.42	.52	.63	.27
Miss	.23	.17	.12	.15
False alarm	.35	.31	.25	.58

Table 2
Mean Correct, Incorrect, and Subjective Details Generated as a Function of Type of Recall Criterion in Experiment 1

Details	Warning		Standard		Forced	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Correct	3.04	1.53	4.38	1.87	7.38	2.69
Incorrect	0.09	0.27	0.48	0.72	0.84	1.59
Subjective	3.54	1.68	4.10	1.59	6.58	2.39

Although the forced condition, on average, generated more incorrect details than did the standard condition, this effect was nonsignificant [$t(118) = 1.59, p = .11$]. Table 2 displays the mean number of details generated across the three description conditions.

Overall description accuracy was also examined by calculating the proportion of correct details over the sum of both correct and incorrect details for each participant. Subjective details could not be judged as either correct or incorrect and thus were left out of this computation. A one-way ANOVA was then used to examine participants' overall description accuracy as a function of the type of recall criterion (forced, standard, or warning recall condition). A significant main effect was observed for the type of recall criterion [$F(2,177) = 6.14, MS_e = 0.014, p < .01, \eta^2 = .07$]. Subsequent analyses indicated that both the standard recall condition ($M = .90, SD = .15$) and the forced recall condition ($M = .91, SD = .13$) generated significantly worse descriptions than did the warning condition ($M = .97, SD = .08$) [$t_s(118) > 3.00, p_s < .01$].

Our second interest was in estimating the effect of participants' description accuracy on subsequent identification of the target face. A discriminant analysis was constructed in which the participants' descriptions (as measured by the number of correct, incorrect, and subjective details generated) were used to predict accuracy on the identification task (hit, miss, or false-alarm response). Two discriminant functions were calculated, with a combined $\chi^2(6) = 21.32, p < .01$. After removal of the first function, there was no significant association between the groups and predictors [$\chi^2(2) = 1.94, n.s.$], indicating that the second discriminant function was not statistically reliable. The first function accounted for 91% of the between-group variability and maximally separated hit responses (group centroid = $-.35$) on the identification task from miss and false-alarm responses (group centroids = $.14$ and $.38$, respectively). The loading matrix of correlations between predictors and this discriminant function suggested that the best predictor of identification accuracy (hit vs. miss or false-alarm response) was the number of incorrect details generated by the participant (factor loading = $.93$). As displayed in Table 3, participants who correctly identified the target generated significantly fewer incorrect details than did those who failed to make a selection [$t(109) = 2.64, p < .01$] and those who selected an incorrect foil [$t(152) = 4.12, p < .001$]. Additionally, a significant correlation was ob-

served between identification accuracy (hit vs. miss or false-alarm response) and overall description accuracy (proportion of correct details generated by participants) when combined across the various description conditions [$r(180) = .33, p < .001$].

Overall, the results of Experiment 1 suggest that forcing participants to generate a more elaborate description of the target significantly impaired their ability to make an accurate identification from a photo lineup. Conversely, warning participants that they should generate only those aspects of the face that they accurately remember significantly enhanced identification accuracy, relative to a control condition. When examining the standard versus control conditions alone, identification accuracy did not differ whether assessed immediately or following a 30-min delay. Furthermore, we did not replicate the "release" of verbal overshadowing effect noted by Finger and Pezdek (1999): Neither the main effect nor any interactions were found involving the postdescription delay. Qualitative analysis of the descriptions indicated that manipulation of participants' response criterion significantly affected the overall quality of their descriptions. Combining across conditions, participants who generated fewer incorrect features were more likely to accurately identify the target, thereby supporting Schooler and Engstler-Schooler's (1990) initial recoding interference account and the influence of retrieval processes within the verbal overshadowing paradigm.

EXPERIMENT 2

Experiment 1 demonstrated that forcing participants to generate misinformation in their verbal descriptions constituted a potent source of interference when they later attempted to identify the target from a photo lineup. A key question that remains, however, is whether this "self-generated misinformation" is more or less powerful than postevent misinformation that is provided to participants by the experimenter (see Roediger et al., 1996; Roediger, McDermott, & Goff, 1997). As reviewed earlier, reliable postevent misinformation effects have been observed with facial stimuli when the erroneous information is provided in the context of a description (Loftus & Ketcham, 1983). However, a condition in which participants are likely to generate errors themselves has not previously been established within the facial memory domain; thus,

Table 3
Mean Correct, Incorrect, and Subjective Details Generated as a Function of Identification Accuracy in Experiment 1

	Details					
	Correct		Incorrect		Subjective	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Hit	4.52	2.81	0.15	0.54	4.39	2.11
Miss	5.38	3.39	0.52	0.83	5.38	2.73
False alarm	5.28	2.79	0.85	1.44	4.92	2.37

to our knowledge, no experimental test of this hypothesis has yet been made.

Within the postevent misinformation paradigm, instructing participants to examine the source of their memory has produced significant performance improvements (Lindsay & Johnson, 1989). However, within the verbal overshadowing paradigm, such source-monitoring instructions have proven ineffective in increasing subsequent identification accuracy (Dodson et al., 1997). Similar to the findings of Ackil and Zaragoza (1998), self-generated misinformation may create a representation that is too similar to that which was perceived and thereby too difficult to distinguish from the original memory. Conversely, experimenter-provided misinformation is an external source of information that may be more readily identified and disregarded (see Goff & Roediger, 1998; Suen-gas & Johnson, 1988). Thus, we also investigated this apparent dissociation between experimenter-provided and self-generated misinformation in Experiment 2.

Experiment 2 followed the general design of Dodson et al. (1997) in examining the differences in magnitude of interference created by self-generated and experimenter-provided descriptions within the verbal overshadowing paradigm. Although similar to the study conducted by Dodson et al., Experiment 2 went further in explicitly manipulating the misinformation contained in the description by either (1) inducing participants to generate erroneous descriptions (forced recall procedure) or (2) providing participants with descriptions, ostensibly written by another individual, but that contained erroneous details (cf. Loftus & Greene, 1980). In addition, source-monitoring instructions were provided to a subset of participants in each condition immediately prior to the photo lineup identification task. A control condition that was not exposed to any misinformation or source-monitoring instructions was also included to provide a baseline for comparison.

On the basis of prior findings (Dodson et al., 1997; Roediger et al., 1996; Roediger et al., 1997; Roediger et al., 1993), we predicted that participants who generated misinformation themselves would produce significantly more misidentifications, relative to those who were provided misinformation by the experimenter postevent and a control condition. Additionally, an interaction between the type of misinformation and source-monitoring instruction was hypothesized, such that only participants in the experimenter-provided misinformation condition would

show accuracy improvement when instructed to consider the source of their memory for the target stimulus.

Method

Participants. The participants were 150 introductory psychology students (62 males and 88 females) who received course credit for their participation. They were tested in groups ranging in size from 2 to 8.

Materials. The target photo and six-person photo lineups were the same as those used in Experiment 1 (Finger & Pezdek, 1999). Again, two versions of the photo lineup were used, such that the target was presented in Position 3 and Position 4 an equal number of times.³

Design and Procedure. A 2×2 between-subjects design was used to examine the effects of manipulations to the type of misinformation (self-generated vs. experimenter-provided) and source-monitoring instruction (source-monitoring instruction vs. no instruction) presented to the participants in the verbal overshadowing paradigm. A control condition was also implemented in which neither independent variable was manipulated. Thirty participants were represented in each cell of the design.

The same procedure was followed as was used in Experiment 1. During the description phase, the participants assigned to the self-generated misinformation condition were given a response sheet with lines numbered from 1 to 25 and were given the forced recall instructions. Those assigned to the experimenter-provided misinformation condition were presented a description that was ostensibly generated by another participant who had previously viewed the same slide. The participants were instructed that they should read over the description for the next few minutes, since they would be asked questions about it at a later time. The description contained 12 facial descriptors and was similar to that created for the participants who served as mock witnesses when conducting lineup fairness analyses in Experiment 1 (see note 1). However, and most important, two aspects of the description were altered, such that they were inconsistent with the target but consistent with other members of the photo lineup (foils). Namely, the hair color reported in the description was altered from brown to auburn, and the complexion of the target was changed from clear to rough looking. The participants assigned to the control condition did not perform either of the misinformation tasks but rather were asked to list as many of the American states as they could recall. The participants in each condition were given 5 min to complete their description or listing task.

Immediately following the misinformation task, the participants were presented with a target-present photo lineup and were given instructions identical to those in Experiment 1. Half of the participants in both the self-generated condition and the experimenter-provided misinformation condition also received source-monitoring instructions with the photo lineup that warned them to consider only the original slide they had viewed and not the description they had previously generated or read. Source-monitoring instructions were identical to those used by Dodson et al. (1997).

Table 4
Proportion Accuracy as a Function of Source Monitoring Instruction and Type of Misinformation in Experiment 2

	No Source Monitoring			Source Monitoring	
	Control	Self-Generated	Experimenter-Provided	Self-Generated	Experimenter-Provided
Hit	.57	.27	.30	.23	.47
Miss	.13	.17	.20	.20	.17
False alarm	.30	.56	.50	.57	.36

Results and Discussion

Identification accuracy. A $2 \times 2 \times 3$ HILOG was performed to examine the influence of type of misinformation (self-generated or experimenter-provided) and source-monitoring instruction (source-monitoring instruction or no instruction) on participants' identification accuracy (hit, miss, or false-alarm response). Results indicated that, although the predicted pattern emerged (see Table 4), the misinformation \times instruction \times accuracy interaction was nonsignificant [$\chi^2(2) = 1.29$, n.s.]. In addition, neither the misinformation \times accuracy [$\chi^2(2) = 2.57$, n.s.] nor the instruction \times accuracy [$\chi^2(2) = 0.90$, n.s.] interaction was significant.

Further planned comparisons examined whether identification performance in the misinformation conditions differed from that in the control condition when the presence of source-monitoring instructions was manipulated. (Note that the control condition was not included in the previous HILOG analysis.) Chi-square analyses on identification performance indicated that participants in the self-generated misinformation condition performed significantly worse than participants in the control condition when source-monitoring instructions were not provided [$\chi^2(2) = 5.93$, $p < .05$]. In addition, participants who were provided misinformation by the experimenter, but not source-monitoring instructions, also tended to perform worse than participants in the control condition, although this effect was not significant [$\chi^2(2) = 4.60$, $p < .10$]. When the participants were warned to consider the source of their memory for the target face, performance on the identification task remained significantly worse for participants in the self-generated misinformation condition [$\chi^2(2) = 7.03$, $p < .05$] but improved to the level of the control condition for participants in the experimenter-provided condition [$\chi^2(2) = 0.60$, n.s.].

As in Experiment 1, the various conditions were similar on the proportion of misses ("not present" responses) [$\chi^2(2) = 0.42$, n.s.] ($M = .17$, $SE = .03$), indicating that any differences in performance were due to variations in the proportion of hits and false alarms. Thus, performance on the identification task was a function of better or worse memory due to manipulations in the recall task, and not a bias against responding to the photo lineup. Additionally, the number of incorrect features generated/read in the two misinformation conditions were rather comparable (self-generated condition, $M = 1.48$, $SD = 1.34$; experimenter-provided condition, 2 incorrect features were included in each description).

Description quality. Experiment 1 demonstrated that participants who generated errors in their verbal descriptions were more likely to misremember the target on a subsequent identification task. In attempting to replicate this result, participants' descriptions in the self-generated misinformation condition were again coded for the number of correct, incorrect, and subjective details. Coders were trained to a criterion of 85% agreement, after which they independently coded each of the 60 descriptions. Interrater reliabilities were sufficient (correct details, $r =$

.92; incorrect details, $r = .91$; subjective details, $r = .92$). For each participant, an estimate was obtained by averaging the coders' ratings within each category.

As in Experiment 1, a discriminant analysis was constructed in which participants' descriptions (as measured by the number of correct, incorrect, and subjective details generated) were used to predict accuracy on the identification task (hit, miss, or false-alarm response). Two discriminant functions were calculated, with a combined $\chi^2(6) = 15.78$, $p < .05$. After removal of the first function, there was no significant association between the groups and predictors [$\chi^2(2) = 1.89$, $p > .05$], indicating that the second discriminant function was not statistically reliable. The first function accounted for 89% of the between-group variability. Replicating the analysis in Experiment 1, the first function maximally separated hit responses (group centroid = $-.80$) on the identification task from miss and false-alarm responses (group centroids = $.75$ and $.11$, respectively). The loading matrix of correlations between predictors and this discriminant function suggested that the best predictor of identification accuracy (hit vs. miss or false-alarm response) was the number of incorrect details generated by the participant (factor loading = $.99$). As displayed in Table 5, participants who correctly identified the target generated significantly fewer incorrect details than did those who failed to make a selection [$t(24) = 3.76$, $p < .001$] and those who selected an incorrect foil [$t(47) = 3.44$, $p < .001$]. Likewise, a significant correlation was observed between identification accuracy (hit vs. miss or false-alarm response) and overall description accuracy (proportion of correct details generated by participants) [$r(60) = .38$, $p < .01$].

Overall, Experiment 2 demonstrated that, relative to a control condition that was not exposed to variations in the type of misinformation or source-monitoring instructions, self-generated and experimenter-provided misinformation both produced interference when participants later attempted to identify the target from a photo lineup. Participants who were both provided a description of the target and instructed to consider the source of their memory showed improvement on the identification task to the level of performance of participants in the control condition. However, such instructions did not assist participants who had previously been forced to generate many details in their own description of the target person, thereby replicating the results of Dodson et al. (1997). As in Experiment 1, a robust verbal overshadowing effect was

Table 5
Mean Correct, Incorrect, and Subjective Details Generated as a Function of Identification Accuracy in Experiment 2

	Details					
	Correct		Incorrect		Subjective	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Hit	7.38	2.78	0.50	0.65	6.53	3.57
Miss	8.05	3.31	2.32	1.72	6.55	2.45
False alarm	7.37	2.43	1.63	1.20	5.52	3.19

found when participants were forced to generate a more elaborate description of the target photo. Furthermore, qualitative analysis of the descriptions produced indicated that generating incorrect details of the target face was predictive of later misidentification in a photo lineup.

GENERAL DISCUSSION

In two experiments, our results indicated that forcing participants to generate a more elaborate description of a target person (forced recall) later impaired their memory on an identification task, both immediately and 30 min following their verbalization. Conversely, warning participants to generate only correct descriptors and not to guess (warning condition) improved their memory in both the immediate condition and the delay condition, relative to the control and forced recall conditions. Moreover, when the participants were supplied with a description that contained erroneous information about the target, subsequent identification accuracy was impaired to the level of those requested to generate a more elaborate description (forced recall). When instructed to consider the source of their memory prior to the identification task, participants who were provided a description by the experimenter improved their identification accuracy to the level of a no-description control condition. In contrast, the identification performance of participants who had self-generated descriptors was not enhanced. Further analysis of the descriptions produced by the participants in both experiments consistently indicated that, across the various recall criterion conditions, erroneous details were most likely to be associated with later misidentification of the target.

Implications for the Verbal Overshadowing Effect

As reviewed earlier, verbal description of faces can enhance, impair, or show no effect on later identification. Although the verbal overshadowing effect has been observed in various studies involving memory for faces, others' attempts at replicating the effect have not always been successful. In the present investigation, we sought to determine whether variations in the response criterion established during the description task would influence the quality of information that is generated by participants, thus further influencing subsequent accuracy on the identification task. Although consistent with Schooler and Engstler-Schooler's (1990) original recoding interference account of the effect, this hypothesis stands in contrast to that of other verbal overshadowing studies (see Schooler et al., 1997), which posited that the simple act of verbalization was responsible for the impairment at recognition. Such a prediction is also consistent with the suppositions of Ericsson and Simon (1993) regarding the procedures to be employed when striving for valid retrospective reports.

Taken together, the results of the present investigation supported these predictions. Failures to replicate the verbal overshadowing effect may have involved differences

in the response criterion (set perhaps via variations in instructions or degree of exhortation used by experimenters) that participants used when attempting to describe the visual stimulus. Inducing participants to lower their response criterion reduced the overall quality of their descriptions, whereas inducing participants to raise their response criterion increased the overall quality of their description. Subsequent attempts at identifying the target were significantly related to these variations in description accuracy. Interestingly, participants in the standard (free response) description condition performed no differently than the no-description controls. As in previous studies, the standard description condition was predicted to exhibit the verbal overshadowing effect; however, our results join with the results of others who have demonstrated a failure to replicate this general effect with standard instructions.

Self-Generated Misinformation

As noted earlier, others have suggested that self-generating misinformation may have a more potent influence on participants' memory for the target stimulus, relative to misinformation that is provided by the experimenter postevent (Ackil & Zaragoza, 1998; Dodson et al., 1997; Roediger et al., 1996; Roediger et al., 1993). Although conclusive evidence for such an effect was not obtained in the present investigation (Experiment 2), two important aspects of self-generated misinformation were identified. First, self-generating erroneous details of a target stimulus produced a significant disruption in memory that was present even after a 30-min delay. Thus, self-generated misinformation did not appear to act as other forms of retroactive interference (or retrieval inhibition) that have been shown to dissipate following moderate delays (Finger & Pezdek, 1999; Wheeler, 1995; see also "spontaneous recovery" discussed in Roediger et al., 1997).

Second, after self-generating misinformation in an initial recall task, participants appeared unable to discriminate between the information they had generated during the description task and those accurate characteristics of the target stimulus they had initially encoded (despite explicit instructions to monitor the source of their memory for the target). Much like the effects observed by others (Roediger et al. 1996; Schooler et al., 1988), inducing participants to respond incorrectly when recalling a target stimulus may lead to later acceptance of the retrieved memory as a veridical representation. As a result, the participants may have experienced significant source confusion when they later attempted to discriminate between details that were internally generated as correct versus incorrect (for similar effects, see Goff & Roediger, 1998, and Suengas & Johnson, 1988).

In addition, monitoring effectiveness may have played a role in the pervasive nature of self-generated misinformation (Koriat, 1995; Koriat & Goldsmith, 1996). Namely, lowering participants' response criterion on an initial description task (forced recall) may have impaired subsequent identification rates by capitalizing on those par-

participants who either ineffectively monitored or failed to monitor the correctness of their verbal description. Thus, during the identification task, participants may have relied on their description when selecting the target, despite its discrepancy with the original stimulus, thereby leading to an incorrect identification.

Processing Shift or Retrieval-Based Interference?

Schooler et al. (1997) argue that the interference resulting from verbalization stems from the type of memory processes brought about by the paradigm (i.e., configural vs. featural processing of faces; see Diamond & Carey, 1986), a theory they termed *transfer inappropriate retrieval* (TIR). They suggest that a participant's description of a target face should contain only those aspects of the stimulus that are readily verbalizable (i.e., featural information), a process that should reduce access to (or overshadow) those aspects that are not recalled or those that cannot be verbalized (i.e., configural information). As a result of the suppression of crucial visual information, participants should later demonstrate difficulty in matching their (description-based) memory for the target with that of the identification stimulus presented them.

Schooler et al. (1997) discuss a number of findings in support of TIR, including evidence on cross-racial effects (Fallshore & Schooler, 1995) and re-presentation of the target face prior to the identification task (Schooler et al., 1996). In the former case, previous studies have shown that own-race faces are likely to be processed in a configural manner, whereas other-race faces are processed in a featural manner (cf. Rhodes, Brake, Taylor, & Tan, 1989). In line with TIR predictions, Fallshore and Schooler

(1995) found that own-race faces were more likely to demonstrate verbal overshadowing effects than were other-race faces. Schooler et al. (1997) also note that TIR is able to account for other findings in the overshadowing literature, such as the discovery that describing one face can interfere with the subsequent identification of a different face (Dodson et al., 1997). In this specific case, Schooler et al. (1997) argue that TIR directly predicts the phenomenon by proposing that describing a face (whether the same as or different from the target) forces participants to engage in verbal-featural processing, which later interferes with the visual-configural processing required at identification.

Within the present set of experiments, our interest has been in investigating the overshadowing effect with respect to the influence of retrieval-based processes on participants' memory for the target face (i.e., recoding interference account; see Schooler & Engstler-Schooler, 1990). We have argued that our results are consistent with the interpretation that manipulation of response criterion on the description task has its most powerful influence on identification by way of its effect on description accuracy. However, Schooler et al. (1997) could contend that inducing a criterion shift during the description task only encouraged our participants to engage in more or less verbal processing of the visual stimulus, thus directly predicting variation on the subsequent identification task.

One manner in which to empirically evaluate these two accounts is illustrated in Figure 1. On the basis of the data from Experiment 1, a path analysis was constructed in which the direct effects of both response criterion and participants' description accuracy on subsequent iden-

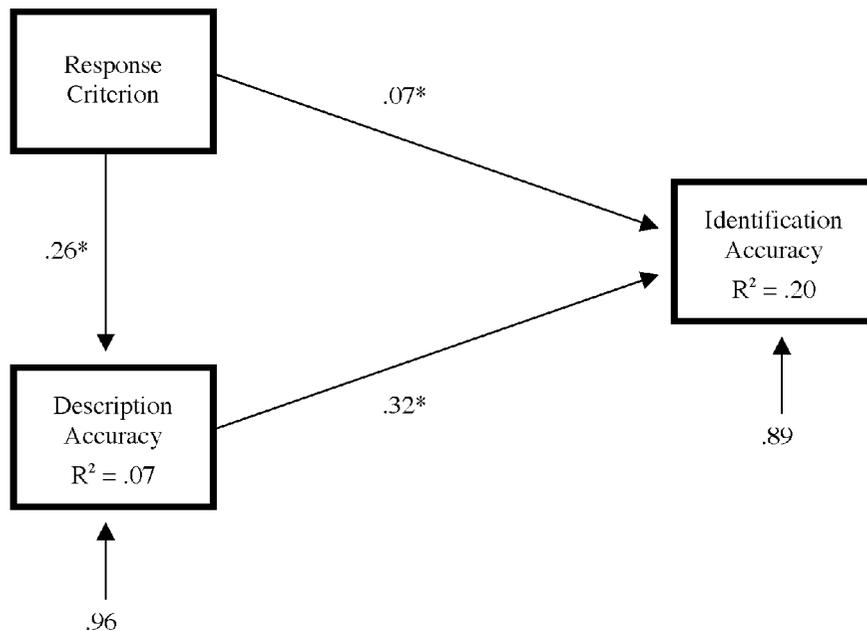


Figure 1. Path analysis of the influence of criterion shifts in the description task on both description accuracy and subsequent identification of a target face in Experiment 1. *bs significant at $p < .05$.

tification were calculated. In addition, the indirect effect of response criterion on participants' description accuracy was examined and controlled for. Results of the path analysis indicated increasing support for the mediated influence of description accuracy on subsequent lineup identification. While the criterion shift of the description task did have a significant direct influence on identification accuracy ($b = .07, p < .05$), it also maintained a highly significant indirect influence on description accuracy ($b = .26, p < .01$). Description accuracy was also strongly related to subsequent identification accuracy ($b = .32, p < .001$). Although neither theoretical account is negated by the present data set, the path analysis suggests that our retrieval-based hypothesis is just as viable as the TIR account, primarily due to the strong mediating influence of description accuracy on later identification of the target face.

Applied Issues

Practically speaking, should law enforcement personnel be informed of the detrimental effects of postexposure verbalization? Witnesses to a crime generally provide a description of the suspect some time following the event itself. However, we know very little of the various techniques or demand characteristics that officers may employ to gain information from a witness. As the present investigation has demonstrated, it is not generally the act of verbalization that later interferes with individuals' ability to identify the target, but rather the retrieval processes and subsequent quality of the description that are most predictive of identification. Hence, as previous studies have shown (see Loftus, 1975), it is important not only that officers refrain from suggesting details to the witness during recall but also that they allow witnesses to establish their own response criterion when providing a description of the event or suspect. Furthermore, it is suggested that demand characteristics be minimized when obtaining a verbal description of a suspect, so that witnesses may actively and effectively monitor the correctness of their responses (including the freedom to withhold information of which they may be unsure).

Conclusions

In closing, the present investigation proposed that the verbal overshadowing effect may be constrained by variation in participants' initial retrieval processes, such that verbalization of a previously viewed stimulus could produce both positive and negative influences on subsequent attempts at recollection. Consistent with this hypothesis, our results indicated that variations in the response criterion that participants adopted in the description task influenced not only the accuracy with which they described the target but also later affected their ability to successfully identify the target from a photo lineup. Furthermore, neither a significant delay nor instructions to monitor the source of their memory improved later identification ac-

curacy for participants who generated errors in their verbal descriptions.

Since only one male target and photo lineup was used, the implications of the present experiments may be limited, due to a lack of stimulus generalization. Regrettably, many other studies in the overshadowing literature also have this shortcoming, largely due to the "trial effect" noted by Schooler et al. (1997) and Schooler et al. (1996).⁴ Future studies should seek not only to address this issue in replicating the present findings but also to create a more active and realistic context such as a live or videotaped crime scenario. These limitations notwithstanding, the present investigation has provided rather clear and compelling evidence for the dramatic influence of retrieval processes in verbal overshadowing. Additionally, it has offered an explanation for some of the inconsistencies found within the overshadowing paradigm and, we hope, may provide a worthwhile framework for future research on factors influencing eyewitness memory.

REFERENCES

- ACKIL, J. K., & ZARAGOZA, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology, 34*, 1358-1372.
- BRIGHAM, J. C., & CAIRNS, D. L. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology, 18*, 1394-1410.
- 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.
- BROWN, E., DEFFENBACHER, K., & STURGILL, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology, 62*, 311-318.
- CARMICHAEL, L., HOGAN, H. P., & WALTER, A. A. (1932). An experimental study of the effect of language on the reproduction of visually perceived forms. *Journal of Experimental Psychology, 15*, 73-86.
- CHANCE, J., & GOLDSTEIN, A. G. (1976). Recognition of faces and verbal labels. *Bulletin of the Psychonomic Society, 7*, 384-386.
- DIAMOND, R., & CAREY, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*, 107-117.
- DODSON, C. S., JOHNSON, M. K., & SCHOOLER, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition, 25*(2), 129-139.
- ERICSSON, K. A., & SIMON, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- FALLSHORE, M., & SCHOOLER, J. W. (1995). The verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 1608-1623.
- FINGER, K., & PEZDEK, K. (1999). The effect of cognitive interview on face identification accuracy: Release from verbal overshadowing. *Journal of Applied Psychology, 84*, 340-348.
- FISHER, R. P., & GEISELMAN, R. E. (1992). *Memory-enhancing techniques for investigative interviewing*. Springfield, IL: Thomas.
- FISHER, R. P., QUIGLEY, K. L., BROCK, P., CHIN, D., & CUTLER, B. L. (1990). *The effectiveness of the cognitive interview in description and identification tasks*. Paper presented at the American Psychology-Law Society, Williamsburg, VA.
- GOFF, L. M., & ROEDIGER, H. L., III (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition, 26*, 20-33.
- GORENSTEIN, G., & ELLSWORTH, P. (1981). Effect of choosing an in-

- correct photograph on a later identification by an eyewitness. *Journal of Applied Psychology*, **65**, 616-622.
- GWYER, P., & CLIFFORD, B. R. (1997). The effects of the cognitive interview on recall, identification, confidence and the confidence/accuracy relationship. *Applied Cognitive Psychology*, **11**, 121-145.
- HERMAN, D. T., LAWLESS, R. H., & MARSHALL, R. W. (1957). Variables in the effect of language on the reproduction of visually perceived forms. *Perceptual & Motor Skills*, **7**, 171-186.
- HOUSER, T., FIORE, S. M., & SCHOOLER, J. W. (1998). *Verbal overshadowing of music memory: What happens when you describe that tune?* Unpublished manuscript.
- KORIAT, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, **124**, 311-333.
- KORIAT, A., & GOLDSMITH, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, **103**, 490-517.
- LINDSAY, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1077-1083.
- LINDSAY, D. S., & JOHNSON, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, **17**, 349-358.
- LOFTUS, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, **7**, 560-572.
- LOFTUS, E. F., & GREENE, E. (1980). Warning: Even memory for faces may be contagious. *Law & Human Behavior*, **4**, 323-334.
- LOFTUS, E. F., & KETCHAM, K. E. (1983). The malleability of eyewitness accounts. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 159-171). New York: Wiley.
- LOVETT, S. B., SMALL, M. Y., & ENGSTROM, S. A. (1992, November). *The verbal overshadowing effect: Now you see it, now you don't*. Paper presented at the annual meeting of the Psychonomic Society, St. Louis.
- MAULDIN, M. A., & LAUGHERY, K. R. (1981). Composite production effects on subsequent facial recognition. *Journal of Applied Psychology*, **66**, 351-357.
- MCKELVIE, S. J. (1976). The effects of verbal labeling on recognition memory for schematic faces. *Quarterly Journal of Experimental Psychology*, **28**, 459-474.
- MELCHER, J. M., & SCHOOLER, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory & Language*, **35**, 231-245.
- READ, J. D. (1979). Rehearsal and recognition of human faces. *American Journal of Psychology*, **92**, 71-85.
- READ, J. D., & SCHOOLER, J. W. (1994, August). *Verbalization decrements in long-term person identification*. Paper presented at the Third Practical Aspects of Memory Conference, College Park, MD.
- RHODES, G., BRAKE, S., TAYLOR, K., & TAN, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, **80**, 313-331.
- RILEY, D. A. (1962). Memory for form. In L. Postman (Ed.), *Psychology in the making* (pp. 402-465). New York: Knopf.
- ROEDIGER, H. L., III, & GUYNN, M. J. (1996). Retrieval processes. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 197-231). San Diego: Academic Press.
- ROEDIGER, H. L., III, JACOBY, J. D., & McDERMOTT, K. B. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory & Language*, **35**, 300-318.
- ROEDIGER, H. L., III, McDERMOTT, K. B., & GOFF, L. M. (1997). Recovery of true and false memories: Paradoxical effects of repeated testing. In M. A. Conway (Ed.), *Recovered memories and false memories* (pp. 118-149). Oxford: Oxford University Press.
- ROEDIGER, H. L., III, WHEELER, M. A., & RAJARAM, S. (1993). Remembering, knowing, and reconstructing the past. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 97-134). San Diego: Academic Press.
- RYAN, R. S. (1992). *Does transfer appropriate processing mediate the verbal overshadowing effect?* Unpublished masters thesis, University of Pittsburgh.
- RYAN, R. S., & SCHOOLER, J. W. (1998). Whom do words hurt?: Individual differences in susceptibility to verbal overshadowing. *Applied Cognitive Psychology*, **12**, S105-S125.
- SCHOOLER, J. W., & ENGSTLER-SCHOOLER, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, **22**, 36-71.
- SCHOOLER, J. W., FIORE, S. M., & BRANDIMONTE, M. A. (1997). At a loss from words: Verbal overshadowing of perceptual memories. In D. Medin (Ed.), *The psychology of learning and motivation* (Vol. 37, pp. 293-334). San Diego: Academic Press.
- SCHOOLER, J. W., FOSTER, R. A., & LOFTUS, E. F. (1988). Some deleterious consequences of the act of recollection. *Memory & Cognition*, **16**, 243-251.
- SCHOOLER, J. W., RYAN, R. S., & REDER, L. M. (1996). The costs and benefits of verbalization. In D. Herrmann, M. Johnson, C. McEvoy, C. Hertzog, & P. Hertel (Eds.), *Basic and applied memory: New findings* (pp. 51-65). Hillsdale, NJ: Erlbaum.
- SUENGAS, A. G., & JOHNSON, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General*, **117**, 377-389.
- WHEELER, M. A. (1995). Improvement in recall over time without repeated testing: Spontaneous recovery revisited. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 173-184.
- WOGALTER, M. S. (1991). Effects of post-exposure description and imaging on subsequent face recognition performance. *Proceedings of the Human Factors Society*, **35**, 575-579.
- WOGALTER, M. S. (1996). Describing faces from memory: Accuracy and effects on subsequent recognition performance. *Proceedings of the Human Factors & Ergonomics Society*, **40**, 536-540.
- YU, C. J., & GEISELMAN, R. E. (1993). Effects of constructing identikit composites on photospread identification performance. *Criminal Justice & Behavior*, **20**, 280-292.

NOTES

1. The fairness of the photo lineup was assessed by having a subset of participants ($n = 10$) create a description of the target person using a description checklist. Modal descriptors were computed and formed into a single description, which was presented to a second subset of participants ($n = 18$), who were asked to select the target from the photo lineup solely on the basis of the description. Lineup fairness parameters were computed by estimating the number of selections across lineup members, with the fairness assumption that selection should be spread evenly across lineup members. According to the scale recommended by Brigham, Meissner, and Wasserman (1999), an overall fairness index = 4 was obtained, indicating no significant bias (functional size = 4.5) or size (effective size = 5.0) considerations.

2. Variation in the placement of the target within the photo lineup (i.e., Position 3 vs. Position 4) did not statistically influence the selection rates of the participants [$\chi^2(2) = 1.84$, n.s.]. Thus, the participants' identification responses across the two lineups were pooled in subsequent analyses.

3. Again, variation in the placement of the target within the photo lineup (i.e., Position 3 vs. Position 4) did not statistically influence the selection rates of the participants [$\chi^2(2) = 0.80$, n.s.]. Thus, the participants' identification responses across the two lineups were pooled in subsequent analyses.

4. Rather unfortunately, the verbal overshadowing paradigm has encountered attenuation of the effect when participants are asked to repeatedly view, describe, and identify a series of faces (see Fallshore & Schooler, 1995; Schooler et al., 1996). Given that we have no current explanation for this loss of the effect across trials (within subjects), studies in the verbal overshadowing paradigm, including our own, have consistently used only a single-item test for each participant.