

## Pictorial encoding reduces false recognition of semantic associates

LANA ISRAEL and DANIEL L. SCHACTER  
*Harvard University, Cambridge, Massachusetts*

High levels of false recognition can be observed after people study lists of semantic associates that all converge on a nonpresented lure word. To test the idea that encoding distinctive perceptual information would help to reduce false recognition, we presented a line drawing representing each associated word during study list presentation and later tested recognition of studied words and lure words. Two experiments revealed marked reductions in false recognition after pictorial encoding, relative after to word encoding. Results suggest that people reject related and unrelated lures because these items lack the distinctive qualities associated with remembered pictures.

False recognition occurs when people incorrectly claim that they have previously encountered a novel test item that is in some way related to a studied item (Underwood, 1965). Although the magnitude of such effects is usually modest, Roediger and McDermott (1995) recently demonstrated exceptionally high levels of false recognition. Modifying a paradigm used by Deese (1959) to induce intrusions on a free recall test, Roediger and McDermott exposed participants to lists of semantic associates that converge on a nonpresented lure word. They reported that false-alarm rates to the critical lure words were comparable to hit rates for studied items, exceeding 80% in some conditions. Subsequent studies have replicated and extended Roediger and McDermott's observations with college students (Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996; Robinson & Roediger, 1997) and used variants of their paradigm to examine false recognition in elderly adults (Norman & Schacter, 1997) and amnesic patients (Schacter, Verfaellie, & Anes, 1997; Schacter, Verfaellie, & Pradere, 1996; for reviews of relevant studies, see Roediger, McDermott, & Robinson, in press, and Schacter, Norman, & Koutstaal, 1998).

One reason for the high false-recognition rates in the Deese/Roediger-McDermott paradigm is that presentation of numerous strong associates may emphasize common semantic features of studied words more than perceptual details of particular items (Schacter et al., 1998). Mather et al. (1997) and Norman and Schacter (1997) found that both true and false recognition relied heavily

on memory for semantic associations and relatively less on memory for distinctive perceptual details; moreover, false recognition was associated with less access to item-specific perceptual information than was true recognition. Using a different paradigm, Brainerd, Reyna, and Kneer (1995) have shown that, when information about specific studied words is made easily accessible during recognition testing, false alarms to semantically related words can be reduced to lower levels than false alarms to unrelated words. Norman and Schacter (1997) hypothesized that false recognition could also be reduced by creating study conditions in which participants are more likely to encode distinctive perceptual details of individual items. To investigate this possibility, we examined whether presenting pictures corresponding to each word in an associative list would produce lower rates of false recognition than would standard study conditions in which pictures are not presented.

Picture-superiority effects have been documented repeatedly on explicit memory tests (for review and discussion, see Paivio, 1971, and Weldon & Roediger, 1987). In contrast to studying words, after studying pictures and presumably forming distinctive perceptual representations, participants should be more likely to demand a specific recollection of a test item before calling it "old." Thus, we hypothesize that participants should make fewer false alarms to related lures after studying lists of semantic associates presented with pictures, relative to the standard paradigm in which only words are presented (cf. Koutstaal & Schacter, in press).

To test this hypothesis, we contrasted true and false recognition in a modified Deese/Roediger-McDermott paradigm under study conditions in which individuals were presented simultaneously with (1) a picture and its auditory label (e.g., participants saw a picture of thread and heard the word *thread*) or (2) a word and its auditory label (e.g., participants saw and heard the word *thread*). Auditory word presentation at study ensured that, for participants in the picture condition, there was no ambiguity concerning the identity of the word depicted by a

---

These experiments are based on a senior thesis carried out by the first author under the supervision of the second author and were supported by National Institute on Aging Grant AG08441 and the Harvard College Research Fund. The authors thank Wilma Koutstaal and Ken Norman, Michele Leichtman, and Alfonso Caramazza for comments and suggestions, Mark Arimoto, Weston Taichiro Eguhi, and Jason Ng for help with preparation of pictorial materials, and Lissa Galluccio for help with preparation of the manuscript. Correspondence should be addressed to D. L. Schacter, Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138 (e-mail: dls@wjh.harvard.edu).

picture. On the recognition test, items were presented in two ways: (1) visual and auditory presentation that reinstated encoding conditions (e.g., participants in the picture encoding group both saw a picture of thread and heard the word *thread*, and participants in the word encoding group both saw and heard the word *thread*), and (2) auditory presentation alone (e.g., all participants heard the word *thread*). The purpose of this manipulation was to determine whether visual reinstatement of pictures or words at test provides perceptual information that could be used to increase the accuracy of recollective judgments.

We performed two experiments that test our hypothesis that studying pictures will reduce false recognition of related lures. Because the experiments are in most respects similar, we report them together. Experiment 1 employed an auditory control condition that was not deemed necessary for Experiment 2, and several details of study and test lists were altered between the two experiments in an attempt to increase the magnitude of the false-recognition effect in Experiment 2 (see Method).

## METHOD

### Participants

Thirty-eight Harvard University undergraduates participated in Experiment 1, and an additional 37 participated in Experiment 2. Two students in Experiment 1 and 1 student in Experiment 2 who did not conform to experimental procedures were excluded, and data from the remaining 36 participants in each experiment were analyzed. In Experiment 1, 12 participants were included in each of three main conditions (word encoding, picture encoding, and auditory control); in Experiment 2, 18 participants were included in each of two main conditions (word encoding and picture encoding). The participants were paid \$8.

### Materials and Design

For Experiment 1, 21 twelve-item study lists were created using the Russell and Jenkins (1954) word association norms and adapting some of the lists used by Roediger and McDermott (1995). Study lists were constructed by selecting the 12 highest associates that could be represented pictorially. The 21 lists were divided into three sets for counterbalancing purposes. Each set was used equally often in each between-subject experimental condition. The participants studied 14 lists, and they were given a 98-item recognition test. The test contained (1) 42 studied items, or *true targets* (items drawn from the 1st, 5th, and 7th list positions of each of the 14 studied lists), (2) 18 new unrelated lures, or *true target controls* (items drawn from the 1st, 5th, and 7th list positions of each of the 7 nonstudied lists; due to an experimental oversight, for 3 of the nonstudied lists only 2 items were included on the recognition test), (3) 28 related lures, or *false targets* (for each of the 14 studied lists, the related lure on which all studied items semantically converge, plus the second highest associate), and (4) 10 new unrelated lures, or *false target controls* (the related lure on which all studied items semantically converge, plus the second highest associate; these 2 items were presented for each of the 7 nonstudied lists, but, due to the oversight noted above, for 4 of the nonstudied lists only 1 new unrelated item was included on the recognition test). Thus, only the true targets were presented at study. The same items appeared equally often as true targets and true target controls or as false targets and false target controls.

For Experiment 2, materials from Experiment 1 were altered by replacing the 2 lists that produced the lowest rates of false recognition across study conditions with 2 new study lists that elicit high levels of false recognition, according to norming data gathered by Stadler, Roediger, and McDermott (1997). Lower associates were substituted with higher associates when possible, and the items that had been used as the second false targets in Experiment 1 were restored to their second po-

sition places on the study lists. To hold list length constant across experiments, items presented in List Position 12 in Experiment 1 were not presented in Experiment 2.

As in Experiment 1, the 21 lists were divided into three sets for counterbalancing purposes. The participants studied 14 lists, and they were given a 63-item recognition test. The test consisted of 28 true targets (drawn from the 1st and 7th list positions of each of the 14 studied lists), 14 true target controls (drawn from the 1st and 7th list positions of each of the 7 nonstudied lists), 14 false targets (the related lure on which all studied items semantically converge for each of the 14 studied lists), and 7 false target controls (the lure for each of the 7 nonstudied lists).

For both experiments, items on the recognition test were randomly assigned to a test presentation mode (visual + auditory, or auditory), with no more than 3 items presented consecutively in the same mode. The recognition test was counterbalanced so that (1) each type of item (i.e., true target, true target control, false target, and false target control) was presented equally often in each of the two presentation modes, and (2) each type of item appeared equally often in the first and second half of the test. Furthermore, items taken from the same study list were at least 8 positions apart on the recognition test, and no more than 2 items of the same type appeared consecutively.

The pictorial stimuli were black-and-white line drawings of list items, and they varied in size (ranging from approximately 3 × 3 cm to 17 × 18 cm, with a modal size of approximately 10 × 11 cm). In general, the drawings possessed similar amounts of detail and novelty, although these features were not systematically controlled. Pictures were scanned on a PowerMacintosh 7600/132 using VistaScan and a UMAX Vista-S6E scanner. Auditory stimuli were recorded on a Macintosh Quadra 150 using SoundEdit Pro. Word stimuli were presented in uppercase in 55-point Geneva typefont. All stimuli were presented on a PowerComputing PowerCenter 132 using PsyScope 1.2b2. The participants heard auditory stimuli through headphones.

The main design for both Experiments 1 and 2 consisted of one between-group variable: (encoding condition [word vs. picture]) and two within-group variables (test presentation mode [visual + auditory vs. auditory] and item type [true target, true target control, false target, and false target control]). In addition, in Experiment 1, an auditory control group was also included in which all stimuli were presented and tested auditorily.

### Procedure

The procedure was identical for the main conditions in the two experiments. The participants were tested individually. They were told that 14 lists of 12 items each would be presented, and that each item was composed of an auditory and visual component. The participants were instructed to pay careful attention to both parts of the item because they would be tested on the items later. Additionally, the participants were told that they would have 1 min to work on a puzzle after presentation of each study list, and that a beep would sound before presentation of the next study list (puzzles included a math subtraction task, a math addition task, and mazes). The visual component of each study item was displayed for 1.5 sec; all auditory components were presented simultaneously in a female voice. Approximately 1.5 sec elapsed between each study item. Presentation of each list took approximately 40 sec. Following presentation of all 14 lists, the participants had 3 min to work on mazes.

After this filler task, the participants received instructions for the recognition test. The participants were asked to indicate whether each item was "old" (i.e., had appeared on one of the study lists) or "new" (i.e., had not appeared on the study lists) by pressing "O" or "N" on the keyboard. They were told to consider both components of an item (i.e., visual and auditory) when making recollective judgments. The participants in the picture condition were also assured that an old picture would never be presented with a new auditory label, nor would a new picture be presented with an old auditory label. Whenever they called a test item "old," the participants made remember/know judgments (Tulving, 1985) using instructions adapted from Rajaram (1993). After the participants completed their judgments, they pressed the space bar for presentation of the next test item. When items were presented in auditory test mode, a cross hair appeared in the center of the computer screen.

These procedures were modified as follows for the auditory control condition of Experiment 1: Statements regarding the dual components

of stimuli were removed from all instructions, and a cross-hair fixation point appeared during the presentation of all study and test items.

**RESULTS**

**True Recognition**

Table 1 presents the proportions of “old” responses and remember/know judgments to true targets, true target controls, false targets, and false target controls as a function of test presentation mode in the word and picture encoding conditions of Experiments 1 and 2.

In both experiments, hit rates were similar in the picture and word encoding conditions. However, the discriminability of old and new items, as indicated by corrected recognition scores that were obtained by subtracting the proportion of “old” responses to true target controls from the proportion of “old” responses to true targets, was markedly higher in the picture encoding condition than in the word encoding condition. An analysis of variance (ANOVA) ( $\alpha = .05$ , for all analyses) on corrected hit rates in Experiment 1 revealed that the main effect for encoding condition (.71 vs. .52 in the picture and word encoding conditions, respectively) was significant [ $F(1,22) = 8.40, MS_e = 0.058, p = .008$ ]. Both the effect for test presentation mode [ $F(1,22) = 3.05, MS_e = .027, p = .095$ ] and the encoding condition  $\times$  test presentation mode interaction [ $F(1,22) = 3.36, MS_e = 0.027, p = .08$ ] approached significance. Further analysis revealed that, in the picture encoding condition, recognition was significantly more accurate when pictures and their auditory labels were presented at test than when only auditory words were presented [.80 vs. .63;  $t(11) = 3.09, p = .01$ ]. By contrast, there was no effect of test presentation mode in the word encoding condition [.52 vs. .52;  $t(11) < 1$ ].

Similar effects were observed in Experiment 2, with corrected recognition rates higher for the participants in the picture (.69) versus the word (.58) encoding condition [ $F(1,34) = 3.90, MS_e = 0.053, p = .053$ ], a significant effect for test presentation mode [ $F(1,34) = 6.90, MS_e =$

$0.028, p = .01$ ], and an encoding condition  $\times$  test presentation mode interaction [ $F(1,34) = 7.13, MS_e = 0.028, p = .01$ ]. As in Experiment 1, these results reflect a significant advantage for corrected recognition rates when items were tested in visual + auditory mode, relative to auditory mode, following picture encoding [.79 vs. .59, respectively;  $t(17) = 5.02, p = .0001$ ], but not following word encoding [.59 vs. .59, respectively;  $t(17) < 1$ ].

Analysis of remember/know judgments revealed that, in both experiments, corrected “remember” responses (i.e., true target “remember” responses minus true target control “remember” responses) were significantly higher in the picture encoding condition than in the word encoding condition [Experiment 1,  $F(1,22) = 13.80, MS_e = 0.053, p < .001$ ; Experiment 2,  $F(1,34) = 5.62, MS_e = 0.057, p = .02$ ]. There were no significant differences in any analysis for corrected “know” responses ( $F_s < 1$ ).

**False Recognition**

As the results in Table 1 show clearly, both experiments revealed a dramatic decrease in “old” responses to non-presented items after studying pictures rather than words only. In Experiment 1, false recognition of false targets was considerably lower in the picture encoding condition (.17) than in the word condition (.47) [ $F(1,22) = 23.75, MS_e = 0.046, p < .0001$ ]. A similar effect was observed for false target controls, with false alarms much less frequent in the picture encoding condition (.04) than in the word encoding condition (.27) [ $F(1,22) = 9.60, MS_e = 0.063, p = .005$ ]. A nearly identical pattern was also observed for true target controls [.05 vs. .21 for picture and word encoding conditions, respectively;  $F(1,22) = 9.65, MS_e = 0.033, p = .005$ ].

The overall false-recognition rate to the false targets was notably lower in Experiment 1 than has been observed in other studies using similar paradigms (e.g., Payne et al., 1996; Roediger & McDermott, 1995; Schacter et al., 1996). This occurred in part because we were constrained to include only those items for which it was possible to generate a picture, but it is probably also attributable to

**Table 1**  
**Proportions of “Old” (O), “Remember” (R), and “Know” (K) Responses on the**  
**Recognition Test as a Function of Item Type, Test Presentation Mode,**  
**and Encoding Condition in Experiments 1 and 2**

Item Type	Test Presentation Mode	Experiment 1						Experiment 2					
		Word Encoding			Picture Encoding			Word Encoding			Picture Encoding		
		O	R	K	O	R	K	O	R	K	O	R	K
True targets		.73	.54	.19	.76	.67	.09	.79	.62	.17	.78	.71	.07
	Auditory	.69	.49	.20	.69	.59	.10	.77	.64	.13	.71	.63	.08
True target controls	Visual+Auditory	.77	.60	.17	.84	.76	.08	.82	.60	.22	.85	.78	.07
		.21	.14	.07	.05	.02	.03	.21	.08	.13	.09	.04	.05
False targets	Auditory	.17	.12	.05	.06	.01	.05	.18	.05	.13	.12	.05	.07
	Visual + Auditory	.25	.15	.10	.04	.03	.01	.23	.10	.13	.06	.02	.03
False target controls		.47	.30	.17	.17	.07	.10	.66	.47	.19	.35	.24	.11
	Auditory	.47	.30	.17	.25	.11	.14	.64	.45	.19	.41	.31	.10
False target controls	Visual + Auditory	.47	.30	.17	.09	.04	.05	.68	.49	.19	.30	.17	.13
		.27	.09	.18	.04	.02	.02	.28	.13	.15	.08	.03	.05
False target controls	Auditory	.20	.05	.15	.05	.02	.03	.26	.12	.14	.08	.03	.05
	Visual + Auditory	.33	.13	.20	.03	.02	.01	.31	.14	.17	.07	.02	.05

the fact that we tested two false targets for each list (the related lure on which all studied items converge, together with another high associate), thus reducing the overall associative strength of the false targets and the study lists.

We restored the second highest associate to the study list in Experiment 2, and Table 1 shows that the overall false-recognition rate was increased substantially. Nonetheless, we observed the same general pattern of results in Experiment 2 as that in Experiment 1: The proportion of "old" responses to false targets was much lower in the picture encoding condition (.35) than in the word encoding (.66) condition; similar patterns were observed for false target controls and true target controls. Separate ANOVAs for each type of lure item revealed significant main effects of encoding condition for false targets [ $F(1,34) = 22.57$ ,  $MS_e = 0.072$ ,  $p < .0001$ ], false target controls [ $F(1,34) = 13.05$ ,  $MS_e = 0.027$ ,  $p = .001$ ], and true target controls [ $F(1,34) = 9.64$ ,  $MS_e = 0.027$ ,  $p = .004$ ].

The effects of encoding condition on false recognition were modified by the effects of test presentation mode. In Experiment 1, an ANOVA on "old" responses to false targets showed a significant effect of test presentation mode [ $F(1,22) = 7.45$ ,  $MS_e = 0.013$ ,  $p = .01$ ] and a significant encoding condition  $\times$  test presentation mode interaction [ $F(1,22) = 5.02$ ,  $MS_e = 0.013$ ,  $p = .036$ ]. The interaction reflects the relatively lower false-recognition rate observed in the visual + auditory test mode, relative to the auditory test mode, following picture encoding (.09 vs. .25, respectively), but not following word encoding (.47 vs. .47, respectively). In Experiment 2, the same general pattern was observed (.30 vs. .41 for visual + auditory and auditory test modes, respectively, after picture encoding, relative to .68 vs. .64, respectively, after word encoding), although the interaction was not significant [ $F(1,34) = 1.54$ ]. In Experiment 1, the proportion of "old" responses to false target controls showed a main effect of test presentation mode [ $F(1,22) = 5.92$ ,  $MS_e = .007$ ,  $p = .02$ ] and an encoding condition  $\times$  test presentation mode interaction [ $F(1,22) = 9.80$ ,  $MS_e = 0.007$ ,  $p = .005$ ]. However, neither of these effects was observed for true target controls in Experiment 1 ( $F_s < 2.14$ ), nor were they significant for false target controls or true target controls in Experiment 2 ( $F_s < 1$ ); so we do not consider them further.

We also analyzed corrected false-recognition scores in which "old" responses to false target controls were subtracted from "old" responses to false targets. In both experiments, there were trends for reduced corrected false recognition in the picture encoding condition, relative to that in the word encoding condition (.13 vs. .20 in Experiment 1; .27 vs. .38 in Experiment 2), but neither effect was significant ( $F_s < 2.53$ ). In both the picture and word encoding conditions of each experiment, the proportion of false alarms to false targets was significantly greater than the proportion of false alarms to false target controls ( $t_s > 4.92$ ,  $p_s < .0005$ ). However, analysis of cor-

rected "remember" false-recognition responses revealed significantly less false recognition following picture encoding (.05) than following word encoding (.21) in Experiment 1 [ $F(1,22) = 10.03$ ,  $MS_e = 0.028$ ,  $p < .005$ ] and a near-significant effect in Experiment 2 [.21 vs. .34;  $F(1,34) = 3.59$ ,  $MS_e = 0.069$ ,  $p = .067$ ]. There were no between-condition differences in corrected "know" false-recognition responses.

### Auditory Control Group

In most previous experiments using the Deese/Roediger-McDermott paradigm, study words were presented auditorily. Experiment 1 included an auditory control condition to determine whether presenting participants with target words both visually and auditorily in the word encoding condition appreciably altered the observed pattern of results. In the auditory control condition, the proportions of "old" responses to true targets (.76), true target controls (.17), false targets (.45), and false target controls (.23) are indistinguishable from corresponding means in the word encoding condition of Experiment 1 ( $t_s < 1$ , for all comparisons).

## DISCUSSION

The two experiments reported here show that when lists of semantic associates are presented with pictures, false alarms to both related and unrelated lures on a subsequent recognition test are sharply reduced, relative to conditions in which only words are studied. Studying pictures enhanced discrimination between old and new items, and it was associated with increases in "remember" responses for studied items (cf. Dewhurst & Conway, 1994) and decreases in "remember" responses for lure items, thus supporting our hypothesis that participants should be able to use the distinctive perceptual information provided by pictures to reduce their susceptibility to false recognition.

This general pattern of results can be usefully conceptualized with reference to the distinction between item-specific or verbatim information and general similarity or gist information (cf. Brainerd et al., 1995; Hintzman & Curran, 1994). As Brainerd et al. contend, item-specific or "verbatim" representations encourage recognition decisions based on identity between the test probe and a previously studied item, whereas gist representations lead to decisions based on similarity. Because words generally lack the specific and varied perceptual information provided by pictures, the participants in the word encoding condition were led to rely on what is salient in the "standard" version of the Deese/Roediger-McDermott paradigm—namely, semantic similarity or gist information.

When semantic similarity serves as the primary basis for one's recognition judgments, individuals are more likely to adopt a liberal recognition criterion, relative to when distinctive perceptual information is available. In our experiments, the participants in the picture encoding condition made fewer false alarms than did the participants in the word encoding condition, despite showing similar hit rates, indicating that they set a stricter criterion for calling an item "old" than did the participants in the word encoding condition. Signal detection analyses were performed on "old" responses to true targets and true target controls, using  $B''_d$  to estimate criterion and  $A'$  to estimate sensitivity (Donaldson, 1996;  $B''_d$  varies between  $-1$  and  $1$ , with higher values indicating more conservative responding, and  $A'$  varies between  $.00$  and  $1.00$ , with higher values indicating greater sensitivity). In both experiments,  $B''_d$  indicated significantly less conservative responding in the word encoding condition (.14 and  $-.02$  in Experiments 1 and 2, respectively) than in the picture encoding condition (.58 and .34; both  $t_s > 2.11$ ,  $p_s < .05$ );  $A'$  indicated significantly lower sensitivity in the word encoding condi-

tion (.83 and .85 in Experiments 1 and 2, respectively) than in the picture encoding condition (.91 and .90; both  $t_s > 2.20$ ,  $p_s < .05$ ; for a more extensive application of this analysis to false recognition data, see Koutstaal & Schacter, in press, and Tussing & Greene, 1997).

Strack and Bless (1994) have claimed that when stimuli are made so memorable that individuals are certain that they should remember studied items, lack of a detailed memory for a nonstudied item provides "diagnostic" evidence that it was not presented. Similar metacognitive processes may have been operative in the picture encoding condition, leading the participants to reject new items that lack the memorable distinctive qualities that are characteristic of studied items.

Our data indicate that reinstating visual information at test, as indicated by comparisons between the visual + auditory and auditory test modes, increased true recognition and reduced false recognition of related lures in the picture encoding condition, but not in the word encoding condition (note, however, that the relevant interaction in Experiment 2 was not significant). We have suggested that participants in the picture encoding condition base their recognition decisions heavily on the availability of perceptual information. Thus, in the picture encoding condition, the additional perceptual information provided through visual reinstatement is highly relevant when making recognition judgments. When a novel picture is presented on the recognition test, the absence of a corresponding perceptual representation of a studied item may be particularly salient, eliciting the kinds of metacognitive processes discussed earlier. However, because participants in the word encoding condition rely less heavily on perceptual information when making recognition judgments, the extra information provided in the visual + auditory test mode is probably not particularly useful or salient.

Although we have emphasized that studying pictures reduces false recognition, it should also be noted that we still observed significant levels of false recognition to related lures following picture encoding. Moreover, analysis of corrected false-recognition scores showed mainly trends for reductions in the picture encoding condition, with the differences achieving significance only for "remember" responses. These relatively modest differences are in line with our suggestion that pictorial encoding reduces false recognition primarily by inducing participants to reject new words—related or unrelated—because they lack the distinctive qualities associated with remembered pictures. Yet the use of such strict criteria was only partially helpful in suppressing the component of the false-recognition effect that is specifically attributable to semantic similarity or gist information. Future studies could examine this finding further by varying word versus picture encoding within subject groups, thereby controlling the between-group criterion differences observed in our experiments. They could also examine whether other kinds of encoding or retrieval manipulations can eliminate entirely a memory illusion that, in our experiments, persisted to some degree under conditions in which people reserved "old" responses for their most distinctive recollections.

## REFERENCES

- BRAINERD, C. J., REYNA, V. F., & KNEER, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language*, **34**, 157-185.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.
- DEWHURST, S. A., & CONWAY, M. A. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1088-1098.
- HINTZMAN, D. L., & CURRAN, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.
- KOUTSTAAL, W., & SCHACTER, D. L. (in press). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory & Language*.
- MATHER, M., HENKEL, L. A., & JOHNSON, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, **25**, 826-837.
- NORMAN, K. A., & SCHACTER, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, **25**, 838-848.
- PAIVIO, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston.
- PAYNE, D. G., ELIE, C. J., BLACKWELL, J. M., & NEUSCHATZ, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory & Language*, **35**, 261-285.
- RAJARAM, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, **21**, 89-102.
- ROBINSON, K., & ROEDIGER, H. L., III (1997). Associative processes in false recall and false recognition. *Psychological Science*, **8**, 231-237.
- ROEDIGER, H. L., III, & MCDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814.
- ROEDIGER, H. L., III, MCDERMOTT, K. B., & ROBINSON, K. (in press). The role of associative processes in creating false memories. In M. A. Conway, S. Gathercole, & C. Cornoldi (Eds.), *Theories of memory II*. Sussex: Psychological Press.
- RUSSELL, W. A., & JENKINS, J. J. (1954). The complete Minnesota norms for responses to 100 words from the Kent-Rosonoff Word Association Test (Tech. Rep. No. 11, Office of Naval Research). Minneapolis: University of Minnesota.
- SCHACTER, D. L., NORMAN, K. A., & KOUTSTAAL, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, **49**, 289-318.
- SCHACTER, D. L., VERFAELLIE, M., & ANES, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology*, **11**, 331-342.
- SCHACTER, D. L., VERFAELLIE, M., & PRADERE, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory & Language*, **35**, 319-334.
- STADLER, M., ROEDIGER, H. L., III, & MCDERMOTT, K. (1997). *Norms for word lists that create false memories*. Manuscript submitted for publication.
- STRACK, F., & BLESS, H. (1994). Memory for nonoccurrence: Metacognitive and propositional strategies. *Journal of Memory & Language*, **33**, 203-217.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychologist*, **26**, 1-12.
- TUSSING, A. A., & GREENE, R. L. (1997). False recognition of associates: How robust is the effect? *Psychonomic Bulletin & Review*, **4**, 572-576.
- UNDERWOOD, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, **70**, 122-129.
- WELDON, M. S., & ROEDIGER, H. L., III (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition*, **15**, 269-280.

(Manuscript received May 7, 1997;  
revision accepted for publication July 7, 1997.)