

“If I had said it I would have remembered it”: Reducing false memories with a distinctiveness heuristic

CHAD S. DODSON and DANIEL L. SCHACTER
Harvard University, Cambridge, Massachusetts

We examined the contributions of decision processes to the rejection of false memories. In two experiments, people studied lists of semantically related words and then completed a recognition test containing studied words, unrelated lure words, and related lure words. People who said words aloud at study were less likely to falsely recognize related lures on the test than were those who heard words at study. We suggest that people who said words at study employed a distinctiveness heuristic during the test whereby they demanded access to distinctive *say* information in order to judge an item as *old*. Even when retrieving *say* information is not perfectly diagnostic of prior study, as in Experiment 2, in which participants both said and heard words at study, people persist in using the distinctiveness heuristic to reduce false memories.

Expectations, metamemorial beliefs, even the format of a memory test can influence the kind and amount of information that is required for making a memory judgment (see, e.g., Dodson & Johnson, 1993, 1996; Lindsay & Johnson, 1989; Marsh & Hicks, 1998; Multhaup, 1995; Schacter, Israel, & Racine, 1999). Theoretically, decision processes are a basic part of both the source-monitoring framework (SMF) of Johnson and colleagues (Johnson, Hashtroudi, & Lindsay, 1993) and the constructive memory framework of Schacter and colleagues (Schacter, Norman, & Koutstaal, 1998). A central assumption of the SMF is that memories are not tagged with an abstract label specifying their origin. Instead, decision processes evaluate the information that is retrieved and attribute it to a source, such as assessing whether enough pictorial information is remembered to respond that an item was seen earlier. Metamemorial beliefs can affect this decision process (Anderson, 1984; Johnson & Raye, 1981). Johnson and Raye noted that people generally “expect self-generated information to have an advantage in memory” (p. 80). A number of studies have shown that people exhibit a bias on memory tests that is consistent with this expectation (e.g., Conway & Gathercole, 1987; Foley, Johnson, & Raye, 1983; Johnson, Raye, Foley, & Foley, 1981; Hashtroudi, Johnson, & Chrosniak, 1989). For instance, Foley et al. (1983) had participants hear words and say words aloud. On a subsequent source memory test, participants were biased to respond that new words, incorrectly judged

old, had been heard earlier rather than said. According to Foley et al., the participants used a test strategy based on their belief about the relative memorability of the study items. Because the participants seemed to expect that they would be more likely to remember said than heard information, test items that were familiar were more often judged to have been heard rather than said.

Schacter et al. (1999; see also Israel & Schacter, 1997) have shown that decision processes can have a powerful impact on the occurrence of false memories. They demonstrated that people can use what they called a *distinctiveness heuristic* to reduce the high false recognition rate of related lure words in a paradigm initially developed by Deese (1959) and recently extended by Roediger and McDermott (1995). In the Deese/Roediger-McDermott (DRM) paradigm, people study lists of words (e.g., *tired*, *dream*) that are related to a nonpresented lure word (e.g., *sleep*). On a subsequent old–new recognition test containing studied words (e.g., *tired*), new unrelated words (e.g., *butter*), and new related lure words (e.g., *sleep*), participants frequently judge that they previously studied the related lure words. However, studying the same items as pictures, instead of words, dramatically reduces false recognition of the related lures. Schacter et al. (1999) argued that rejecting related lure words after picture encoding, as compared with word encoding, stems from participants’ metamemorial belief that they ought to remember this distinctive pictorial information. By this view, participants use a distinctiveness heuristic whereby they demand access to pictorial information as a basis for judging items as previously studied; the *absence* of memory for this distinctive information indicates that the test item is new. In contrast, participants who studied words would not expect detailed recollections about studied items and, thus, would not base recognition decisions on the presence or absence of memory for distinctive information.

This research was supported by National Institute on Aging Grant AG08441 and by a grant from the Human Frontiers Science Program. We appreciate the helpful comments of Reed Hunt, Marcia Johnson, Kathleen McDermott, and John Wixted. Correspondence concerning this article should be addressed to C. S. Dodson, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138 (e-mail: cdodson@wjh.harvard.edu).

Table 1
Proportion of Items Judged *Old* in Experiments 1 and 2

Item Type	Experiment 1 (Between Groups): Encoding Condition		Experiment 2 (Within Groups): Encoding Condition	
	Hear	Say	Hear	Say
Targets	.71	.76	.70	.79
Target controls	.16	.16	.12	.12
False targets	.72	.51	.57	.58
False target controls	.29	.28	.14	.14

Although the studies of Schacter et al. (1999) are consistent with the notion that participants can use a distinctiveness heuristic to reduce false recognition, it is possible that the effects are specific to pictorial stimuli. In the present experiments, the generality of the distinctiveness heuristic as a tool for reducing false recognition is explored. These experiments draw on the findings from Johnson and colleagues, especially Foley et al. (1983), that participants' expectations about the relative memorability of generated and perceived items can produce a response bias on the test. Participants in the say encoding condition of Foley et al.'s study appeared to use a kind of distinctiveness heuristic, inferring that the lack of memory for distinctive *say* information indicated that the test item was earlier heard. We reasoned that a say–hear encoding manipulation within the DRM false memory paradigm might reveal the use of a distinctiveness heuristic. If so, saying words at study should yield fewer false recognition responses to the related lure items than does hearing words at study. Importantly, observing such effects with these purely verbal materials would extend the generality of the distinctiveness heuristic as a tool for reducing false recognition beyond the domain of pictorial encoding initially examined by Schacter et al. (1999).

EXPERIMENT 1

Method

Participants. Forty-two Harvard University undergraduates participated in this experiment. Each was paid \$8.

Materials and Design. We created 24 study lists that were based on the lists used by Roediger and McDermott (1995); each list contained 15 semantically related words. For counterbalancing purposes, the 24 lists were divided into three sets of 8. The participants studied 16 lists and then completed a recognition test containing items from each of the 24 lists. During both study and test phases, all the items appeared centered on an Apple Power Macintosh G3 computer in bold lowercase letters and in 24-point Geneva typefont. For both say and hear encoding groups, all the study words were seen for 1.5 sec, and each word was separated by a 1.5-sec delay. The participants in the hear encoding condition also listened to items at conversational levels through headphones. Approximately 400 msec elapsed between the visual and the subsequent auditory presentations of the word. In the say encoding condition, the items were seen only, although the participants spoke the word aloud upon seeing it.

Following the study phase, the participants completed a 96-item recognition test. The test contained 48 targets or studied words (from Positions 1, 8, and 10 of the 16 studied lists), 24 target controls or unrelated lures (from Positions 1, 8, and 10 of the 8 non-

studied lists), 16 false targets or related lures (the critical lure that is semantically related to the items in each of the 16 studied lists), and 8 false-target controls or unrelated lures (the critical lure that is semantically related to the items in each of the 8 nonstudied lists). The test was constructed so that the items were randomly intermixed and also obeyed the following constraints: (1) Each type of item (target, target control, false target, and false-target control) appeared equally often in each quarter of the test; (2) no more than two items of the same type appeared consecutively; and (3) items from the same study list were at least eight positions apart.

In sum, the main design consisted of one between-groups variable, the say versus hear encoding condition, and one within-groups variable, the type of item.

Procedure. The participants were tested individually and were told to pay attention to the words that they would see because their memory for these words would be tested later. In the hear encoding condition, the participants were told that they would hear each of the words in addition to seeing it. The participants in the say encoding condition were instructed to say aloud each word that they saw. All the participants were told that they would see 16 different lists of words, with 15 words in each list. They were instructed that each list was separated by a 1-min break, during which time they would work on a puzzle. After studying the 16 lists, the participants worked on puzzles for 3 min before receiving the recognition test.

For the recognition test, the participants indicated whether each test item was *old* (i.e., it had been seen earlier on 1 of the 16 study lists) or *new* (i.e., it had not appeared on the study lists) by pushing the “a” or the “;” key, which had been labeled “old” and “new,” respectively. All test words were presented visually in the center of the screen, with the response options “Old” or “New” appearing 2 cm below the word. After each response, the screen cleared and was followed by a 1.5-sec delay before the presentation of the next test item.

Results

Table 1 displays the central finding of the experiment: False targets (related lures) are incorrectly judged *old* much less often in the say than in the hear encoding condition, as was confirmed by an analysis of variance [ANOVA; $F(1,40) = 7.90, p < .01$]. In contrast, there was no difference between the say and the hear encoding conditions in the recognition rate of target items, target control items, and false-target control items [all $F_s(1,40) < 1.43$]. This pattern was confirmed by a 2 (say/hear encoding) \times 2 (target/false target) ANOVA, which yielded significant main effects of encoding condition [$F(1,40) = 5.13, MS_e = 0.041, p < .05$], and item type [$F(1,40) = 4.56, MS_e = 0.043, p < .05$], and, importantly, a significant interaction [$F(1,40) = 5.30, p < .05$]. Recognition rates of targets (.71) and false targets (.72) were nearly identical after hearing words at study [$F(1,40) < 1$], in line with other studies using the DRM paradigm (Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996; Roediger & McDermott, 1995; Schacter, Verfaellie, & Pradere, 1996). After saying words, however, false targets (.51) were less likely to be called *old* than were targets [.76; $F(1,40) = 9.86, p < .01$].

To identify the contributions of sensitivity and response bias to our findings, we performed signal detection analyses, using A' as an estimate of sensitivity and $B''D$ as an estimate of bias (Donaldson, 1992; Snodgrass

Table 2
Values of Sensitivity (A') and Response Bias ($B''D$)
for Items in Experiments 1 and 2

Comparison	Experiment 1 (Between Groups): Encoding Condition		Experiment 2 (Within Groups): Encoding Condition	
	Hear	Say	Hear	Say
	Values of Sensitivity (A')			
Target versus target control	.84	.86	.86	.88
Target versus false target	.50	.69	.62	.68
False target versus false target control	.77	.66	.76	.77
Values of Bias ($B''D$)				
Target versus target control	.40	.33	.54	.45
Target versus false target	-.57	-.37	-.30	-.46
False target versus false target control	.13	.43	.63	.62

& Corwin, 1988).¹ These analyses are based on those used by Koutstaal and Schacter (1997), Schacter, Verfaellie, Anes, and Racine (1998), and Schacter et al. (1999). Values of A' can vary between 0 and 1, with higher values indicating better sensitivity and chance performance corresponding to a value of .5. Scores for the response bias measure, $B''D$, can vary between -1 (extremely liberal responding) and 1 (extremely conservative responding).

Following other studies from our laboratory (Koutstaal & Schacter, 1997; Schacter et al., 1999; Schacter, Verfaellie, et al., 1998), we analyzed the data, as is shown in Table 2, in three different ways. First, the say and hear encoding conditions did not differ in either sensitivity or bias associated with discriminating targets (studied items) from target controls (unrelated lures) [$F_s(1,40) < 1$]. Second, the participants who said words were better able to discriminate targets from false targets ($A' = .69$) than were those participants who heard words, who were at chance in this comparison [$A' = .50$; $F(1,40) = 14.79$, $MS_e = 0.027$, $p < .001$]. In this target/false-target comparison, however, there was generally more conservative responding after say than after hear encoding, although this difference was not significant [$F(1,40) = 2.43$]. Finally, we analyzed the sensitivity and bias scores for the comparison between responses to false targets (related lure words) and false-target controls. Hear encoding produced generally greater A' scores [$F(1,40) = 3.74$, $MS_e = 0.035$, $p = .06$], and more liberal responding [$F(1,40) = 3.21$, $MS_e = 0.280$, $p < .08$], than did say encoding, although both of these effects were only marginally significant.

Discussion

Experiment 1 demonstrates that saying, as compared with hearing, words at encoding reduces the false recognition rate of semantically related lures on the test. These results extend those of Israel and Schacter (1997) and Schacter et al. (1999) by showing that reductions of false recognition produced by distinctive encoding are not restricted to pictorial materials (see also Smith & Hunt, 1998).

There are a number of aspects of our data that suggest that participants use a distinctiveness heuristic to reject false targets. First, suppressing the occurrence of false recognition in the say encoding condition was specific to false targets and was not a by-product of an overall reduction in recognition of all the items: The studied words were recognized at nearly identical rates in the two encoding conditions. Second, the pattern of results from the signal detection analyses paralleled the recognition data. Say and hear encoding yielded comparable levels of sensitivity (i.e., values of A') for distinguishing targets (studied words) from target controls (unrelated lures). However, say encoding produced greater sensitivity than did hear encoding for distinguishing the false targets from the targets. In general, people tended to be more conservative, as indexed by $B''D$, after say than after hear encoding. Overall, we observed a selective reduction in false-recognition responses to the false targets in the say encoding condition.

The previous analyses indicate that the participants used a distinctiveness heuristic after say encoding to reject related lure items. The more conservative responding in the say encoding condition, we argue, is the product of participants' metamemorial beliefs about the kind of information they feel they ought to remember. After saying words at study, the participants employed a heuristic during the test, whereby they demanded access to distinctive say information in order to judge an item as *old*. In other words, the participants' reasoning may be characterized as "If I had said it I would have remembered it." Thus, the absence of memory for saying target items indicated that the word was not encountered during the study phase.

There are two curious aspects to our results, however. First, in both encoding conditions, there was a high false-recognition rate for the false-target controls. This seems to be a typical finding and probably occurs because the false-target controls tend to have higher word frequencies than do the target controls (see, e.g., Arndt & Hirshman, 1998; Israel & Schacter, 1997; Seamon, Luo, &

Gallo, 1998). Second, saying words during encoding produced a subsequent reduction in *old* responses to false targets, but not in *old* responses to other new items. If the distinctiveness heuristic is a strategy that is applied throughout the test, we would expect lower false-recognition rates for all of the new items—the unrelated new items (i.e., the target controls and false-target controls), in addition to the false targets. Indeed, Schacter et al. (1999) used a similar design and found that picture encoding, as compared with word encoding (using a between-groups design), produced lower false-recognition responses to both the false targets and to the true- and false-target controls. We suggest that say encoding yields less salient or distinctive recollections than does picture encoding. Consequently, after saying words aloud, people may apply the distinctiveness heuristic selectively, such as for test items that are very familiar. For these items, subjects may respond *old* only when they can remember information about having said this item at study. Therefore, although the false targets may feel very familiar, the absence of memory for say information is evidence that they were not studied previously. This strong sense of familiarity would not be associated with unrelated new items, and hence, the distinctiveness heuristic may not be needed or used.

However, consider an alternative account of the data that does not require a distinctiveness heuristic. Specifically, suppressing false recognition following say encoding may depend on remembering list-specific information about studied items. For example, when presented with the related lure item *sleep* on the recognition test, the participants in the say encoding condition may have recalled saying related words, such as *tired*, *bed*, *dream*, and so forth. Because they could remember saying these words but could not recall saying *sleep*, the participants may have concluded that *sleep* was never presented during the study phase. This list-specific account hinges on participants' recollecting items from particular study lists and using this list-specific information to avoid false recognition responses. On this account, saying, as compared with hearing, words increases the probability of recollecting list-specific information.

EXPERIMENT 2

One way to evaluate the influence of a distinctiveness heuristic is to construct a situation in which it is difficult to use it. Following the logic of Schacter et al. (1999), we implemented this strategy in Experiment 2 by using a within-groups design, instead of the between-groups design used in Experiment 1. That is, the participants heard some words and said others during the study phase and then completed a recognition test. In this design, the distinctiveness heuristic should be ineffective, because there is no longer a particular kind of information that is solely diagnostic of a test item's oldness or newness. Whereas in Experiment 1 the absence of information about having

said an item suggested that the item was new, this was not the case in Experiment 2. Therefore, if suppressing false recognition is primarily attributable to the distinctiveness heuristic, a within-groups manipulation of say versus hear encoding should eliminate the suppression effect. On the other hand, if false-recognition suppression is based on recalling list-specific information for the spoken lists and not for the heard lists, we should observe a selective suppression effect for the spoken lists with the within-groups manipulation.

Method

Participants. The participants were 27 Harvard University undergraduates, who were each paid \$8.

Materials and Design. The study and test materials were identical to those used in Experiment 1. The only difference was that each participant said eight of the studied lists and heard eight of the lists. No more than two lists were studied consecutively in the same manner.

Procedure. The one difference from Experiment 1 was that the participants were instructed that they would say aloud some of the lists of words and hear other lists of words. Before each list, instructions appeared on the screen indicating whether the list was to be spoken aloud or heard.

Results

As can be seen in Table 1, hit rates were higher for spoken words (i.e., targets) than for heard words. Importantly, false-recognition rates were nearly identical to the false targets related to spoken lists and to heard lists. A 2×2 ANOVA, with factors of encoding (say vs. hear) and item type (target vs. false target), yielded a marginally significant effect of encoding [$F(1,26) = 3.04$, $MS_e = 0.021$, $p < .10$], a significant effect of item-type [$F(1,26) = 13.86$, $MS_e = 0.051$, $p < .01$], and importantly, a significant interaction [$F(1,26) = 4.08$, $MS_e = 0.01$, $p = .05$]. Although the participants showed higher recognition rates of spoken words than of heard words [$F(1,26) = 10.32$, $p < .01$], there were no differences in the false-recognition rate of false targets in the say and the hear encoding conditions [$F(1,26) < 1$].

Table 2 presents the estimates of sensitivity (A') and bias ($B''D$) for our three critical comparisons. An ANOVA of the A' and $B''D$ scores associated with discriminating targets from target controls revealed higher sensitivity scores and more liberal responding for said words than for heard words [$F(1,26) = 9.50$, $MS_e = 0.025$, $p < .01$, and $F(1,26) = 5.38$, $MS_e = 0.021$, $p < .05$, respectively]. For the comparison of the rates at which targets and false targets were called *old*, saying words produced greater discriminability (i.e., higher A' scores) than did hearing words [$F(1,26) = 6.37$, $MS_e = 0.008$, $p < .05$]. However, the participants were more conservative after having heard words than after having said words, although this difference was marginally significant [$F(1,26) = 3.57$, $MS_e = 0.093$, $p = .07$]. Finally, there were nearly identical A' and $B''D$ scores in the two encoding conditions for the comparison between false targets and false-target controls [$F_s(1,26) < 1$].

Discussion

Saying words, as compared with hearing words, produced higher hit rates and higher A' scores associated with discriminating targets from target controls. Despite higher levels of memory for spoken than for heard words, proportions of *old* responses to false targets were nearly identical in these two encoding conditions. The false-recognition rates of the false targets and the A' scores associated with distinguishing between false targets and false-target controls were no different after having said or heard words. These results support our prediction that when the distinctiveness heuristic is rendered ineffective by having people both hear and say words at encoding, there is no difference between the two encoding conditions in the false-recognition rate of the false targets. Moreover, these results are evidence against the alternative list-specific account of the suppression effect in Experiment 1. If say encoding allowed people to recall particular list items (e.g., *tired, bed, dream*, etc.) as a way of ruling out the false target (i.e., *sleep* was not on the list), there should have been a selective reduction in this experiment in the false-recognition rate of the false targets related to the said lists. Since there was no selective reduction, we conclude that when it is, to some degree, diagnostic of an item's oldness, the distinctiveness heuristic is the primary mechanism for suppressing false memories in this paradigm.

However, if having people hear some lists of words and say other lists renders the distinctiveness heuristic ineffective, we would expect the false-recognition rates to the false targets in both encoding conditions in Experiment 2 to rise to the level of the hear encoding condition in Experiment 1—the condition in which people show robust false memories. As can be seen in Table 1, the false-recognition rates to the false targets in both conditions of Experiment 2 are somewhat suppressed and are comparable with the false-recognition rate in the say condition in Experiment 1. Schacter et al. (1999) observed a similar finding (see their Figure 1) across their two experiments. Taken together, these results suggest that people still attempt to use the distinctiveness heuristic even though it is not perfectly diagnostic of an item's *oldness*, and does not produce selective suppression of false alarms for the say encoding condition, as compared with the hear encoding condition. If people are heavily weighting their memories for say information, this would contribute to the higher recognition rates for the said words than for the heard words (whereas there were no differences in the recognition rates of said and heard words in Experiment 1, which was a between-subjects manipulation).

GENERAL DISCUSSION

In our experiments, we investigated the contributions of a distinctiveness heuristic to reducing false memories in the DRM paradigm. In Experiment 1, we found that the participants who said words at study showed less false recognition of related lure words than did the participants

who heard words at study. Saying words aloud—like studying pictures in Schacter et al.'s (1999) study—provides a basis for employing a distinctiveness heuristic during the test. That is, saying words at study leads people to expect to remember information about having said the word; the absence of this expected information suggests that the test word is new. Even when retrieving say information is no longer diagnostic of prior study, as in Experiment 2, in which participants both said and heard words at study, people persist in attempting to use the distinctiveness heuristic to reduce false memories. The absence of a difference between the two encoding conditions in the false-recognition rates of the false targets in Experiment 2 indicates that say encoding does not allow people to remember list-specific information as a way of rejecting false targets, disconfirming an alternative list-specific account of the suppression effect in Experiment 1. In short, the distinctiveness heuristic seems to be the main mechanism underlying false-recognition suppression after picture or speak-aloud encoding.

The distinctiveness heuristic is consistent with one of the basic tenets of the source-monitoring framework of Johnson and colleagues—namely, that people use various decision strategies to evaluate activated information and attribute it to a particular source (Johnson et al., 1993). Hicks and Marsh (1999) proposed an account similar to our distinctiveness heuristic to explain conditions in which they observed reduced levels of false recall of critical lures in the DRM paradigm. They found that people were less likely to recall a critical lure after studying a list of related words from two different sources, such as hearing some words and seeing other words as anagrams, than after studying a list of related words from the same source, such as hearing all of the items. Hicks and Marsh argue that “presenting DRM items from multiple sources can potentially create diagnostic evidence that can be used to edit out false memories” that lack this quality (p. 1206). These studies, together with our findings and those of Johnson and colleagues (e.g., Foley et al., 1983), indicate that an inference process based on the absence of memory for expected distinctive characteristics is a general retrieval strategy that is used during recall, recognition, and source monitoring (for a different perspective on retrieval strategies and criterion setting in the DRM paradigm, see Miller & Wolford, 1999; cf. Roediger & McDermott, 1999; Wixted & Stretch, 2000).

It is worth noting that Smith and Hunt (1998) report lower false recall and recognition of the false targets after visual than after auditory presentation of the study words. Because people saw the words in the say condition of our experiments, it is conceivable that our suppression effect is attributable, at least in part, to the visual presentation of the study words. However, two lines of evidence indicate that this suppression effect is due to saying the words, rather than to visual presentation of the words. First, visual presentation of the study words was held constant across our study conditions. For instance, in Experiment 1, people either heard and saw the words or

said and saw the words. If visual presentation were entirely responsible for the suppression effect, we should have found no difference in the false-recognition rate of the false targets in these two conditions. Second, Israel and Schacter (1997) compared visual and auditory presentation with auditory-only presentation and found no difference in subsequent true- and false-recognition rates. Seeing the words at study did not affect subsequent recognition performance (see also Arndt & Hirshman, 1998; Seamon et al., 1998, who find robust false-recognition rates after visual presentation of the study words). In sum, our experiments indicate that saying words at encoding, as opposed to seeing them, leads people to expect to remember these words at test, thus yielding the use of the distinctiveness heuristic.

A metacognitive heuristic similar to the distinctiveness heuristic has been used to explain the pattern of false-recognition data reported by Strack and Bless (1994). Strack and Bless required participants to study related words, such as different tools, together with a few words that represented unique instances of other categories, such as a book or a shoe. On a subsequent recognition test, the participants exhibited lower false recognition of salient distractor items (e.g., a bouquet of flowers) than of nonsalient distractor items (e.g., a kind of tool). Strack and Bless argued that the salience of the distractor item was diagnostic of its *newness* and, thus, served as a basis for its rejection on the recognition test. However, Rotello (1999) noted that salient distractors are more dissimilar to studied items, and consequently less familiar, than are nonsalient distractors. Therefore, the similarity of the distractor to studied items may explain the pattern of greater rejection rates of salient than of nonsalient distractors. In a series of experiments, Rotello provided evidence in support of this similarity account of the false-recognition data in the Strack and Bless paradigm.

The distinctiveness heuristic discussed here and the heuristic proposed by Strack and Bless (1994) both invoke a decision rule whereby the failure to retrieve particular information about an event can be diagnostic of the event's nonoccurrence. Hence, it is important to ask whether a similarity account, and not the distinctiveness heuristic, accounts for the results in our experiments. A similarity account explains recognition decisions in terms of familiarity, or the overall similarity (reflected by the amount of activation) between a test item and the memorial representation of studied items (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988). Test items that are sufficiently similar to studied items are judged *old*, and those that are not are judged *new*. By this view, the typically large false-recognition rates of related lures in the DRM paradigm occur because of their high similarity to studied items (Arndt & Hirshman, 1998). To explain our false-target rejection results in Experiment 1, this similarity account would contend that say encoding is sufficiently different from hear encoding that the false targets are, in some sense, less similar to previously said words than to previously heard words. Put another way, this

similarity account would have to argue that lower false recognition of the critical lures after say encoding occurs because the critical lures are less familiar (i.e., less similar) after say than after hear encoding. There are problems, however, with this similarity explanation of our results. First, if say and hear encoding produced sufficiently different representations to account for the suppression of false memories in Experiment 1, this similarity explanation would predict differences in the recognition rates of spoken and heard words. This prediction was not supported by the data: In Experiment 1, the said and heard words yielded nearly identical *A'* old–new recognition scores. Second, Experiment 2 provides further evidence against the similarity account. The participants in Experiment 2 showed slightly higher *A'* old–new recognition scores for spoken than for heard words. Nonetheless, there was no difference in the false-recognition rate of false targets that were related to spoken and heard lists. Thus, the similarity account offers no ready explanation of these results. According to the distinctiveness heuristic account, in contrast, the absence of expected say information is diagnostic of a test item's nonoccurrence.

In sum, beliefs and expectations about the workings of memory affect judgments about the origins of past experiences (e.g., Johnson & Raye, 1981). People are aware that different kinds of encoding activities yield memories that vary in strength, detail, and vividness. This metamemorial knowledge can be used as a distinctiveness heuristic to reject false memories.

REFERENCES

- ANDERSON, R. E. (1984). Did I do it or did I only imagine doing it? *Journal of Experimental Psychology: General*, **113**, 594-613.
- ARNDT, J., & HIRSHMAN, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory & Language*, **39**, 371-391.
- CONWAY, M. A., & GATHERCOLE, S. E. (1987). Modality and long-term memory. *Journal of Memory & Language*, **26**, 341-361.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.
- DODSON, C. S., & JOHNSON, M. K. (1993). Rate of false source attributions depends on how questions are asked. *American Journal of Psychology*, **106**, 541-557.
- DODSON, C. S., & JOHNSON, M. K. (1996). Some problems with the process dissociation approach to memory. *Journal of Experimental Psychology: General*, **125**, 181-194.
- DONALDSON, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, **121**, 275-277.
- FOLEY, M. A., JOHNSON, M. K., & RAYE, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development*, **54**, 51-60.
- GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.
- HASHIROUDI, S., JOHNSON, M. K., & CHROSNIAK, L. D. (1989). Aging and source monitoring. *Psychology & Aging*, **4**, 106-112.
- HICKS, J. L., & MARSH, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1195-1209.
- HINTZMAN, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528-551.
- ISRAEL, L., & SCHACTER, D. L. (1997). Pictorial encoding reduces false

- recognition of semantic associates. *Psychonomic Bulletin & Review*, **4**, 577-581.
- JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67-85.
- JOHNSON, M. K., RAYE, C. L., FOLEY, H. J., & FOLEY, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology*, **94**, 37-64.
- KOUTSTAAL, W., & SCHACTER, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory & Language*, **37**, 555-583.
- LINDSAY, D. S., & JOHNSON, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, **17**, 349-358.
- MARSH, R. L., & HICKS, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1137-1151.
- 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.
- MILLER, M. B., & WOLFORD, G. L. (1999). The role of criterion shift in false memory. *Psychological Review*, **106**, 398-405.
- MULTHAUP, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology & Aging*, **10**, 492-497.
- 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.
- 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.
- 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. (1999). False alarms about false memories. *Psychological Review*, **106**, 406-410.
- ROTELLO, C. M. (1999). Metacognition and memory for nonoccurrence. *Memory*, **7**, 43-63.
- SCHACTER, D. L., ISRAEL, L., & RACINE, C. A. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory & Language*, **40**, 1-24.
- 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., & RACINE, C. (1998). When true recognition suppresses false recognition: Evidence from amnesic patients. *Journal of Cognitive Neuroscience*, **10**, 668-679.
- 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.
- SEAMON, J. G., LUO, C. R., & GALLO, D. A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science*, **9**, 20-26.
- SMITH, R. E., & HUNT, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review*, **5**, 710-715.
- SNODGRASS, J. G., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.
- STRACK, F., & BLESS, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory & Language*, **33**, 203-217.
- WIXTED, J. T., & STRETCH, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review*, **107**, 368-376.

NOTE

1. We also analyzed the data from both experiments with d' and C . This produced nearly identical results as the A' and $B''D$ analyses, except for the comparison in Experiment 1 between the false targets and the false-target controls in the two encoding groups. The say and hear encoding groups were marginally significantly different with the A' analysis, but they were not different in the d' analysis. This discrepancy between these two measures does not alter the conclusions of the paper. Therefore, for the sake of comparability with previous papers on this topic, we report the A' scores.

(Manuscript received June 22, 1999;
revision accepted for publication December 20, 1999.)