

Associative false recognition occurs without strategic criterion shifts

DAVID A. GALLO, HENRY L. ROEDIGER III, and KATHLEEN B. McDERMOTT
Washington University, St. Louis, Missouri

In the DRM (Deese/Roediger and McDermott) false memory paradigm, subjects studied lists of words associated with nonpresented critical words. They were tested in one of four instructional conditions. In a standard condition, subjects were not warned about the DRM Effect. In three other conditions, they were told to avoid false recognition of critical words. One group was warned before study of the lists (affecting encoding and retrieval processes), and two groups were warned after study (affecting only retrieval processes). Replicating prior work, the warning before study considerably reduced false recognition. The warning after study also reduced false recognition, but only when critical items had never been studied; when critical items were studied in half the lists so that subjects had to monitor memory for their presence or absence, the warning after study had little effect on false recognition. Because warned subjects were trying to avoid false recognition, the high levels of false recognition in the latter condition cannot be due to strategically guessing that critical test items were studied. False memories in the DRM paradigm are not caused by such liberal criterion shifts.

Roediger and McDermott (1995) reported a memory illusion induced by presenting semantically related materials. On the basis of earlier work by Deese (1959), they devised lists of associates (e.g., *bed, rest, awake*, etc.) of a critical item (*sleep*) that was not studied. Subjects frequently recalled the critical item on an immediate free recall test, and, on recognition tests, the false alarm rate to the critical items was roughly equivalent to the hit rate for list items. Further, subjects frequently claimed to subjectively “remember” the presentation of the critical items in the study list, using Tulving’s (1985) *remember/know* judgment. These observations have now been replicated dozens of times (e.g., McDermott, 1996; Payne, Elie, Blackwell, & Neuschatz, 1996; see Roediger, McDermott & Robinson, 1998, for a partial review). The underlying assumption of this research is that the paradigm (subsequently labeled the Deese/Roediger and McDermott or DRM paradigm) provides a reliable method for creating and studying false memories.

Miller and Wolford (1999) questioned the basis of these associative errors and argued that they were caused by criterion shifts. They proposed that subjects in this task develop meta-knowledge about each list’s theme, or re-

lational structure. At test, subjects are more willing to call a test item “old” (i.e., respond more liberally) if they judge it to be consistent with this meta-knowledge. Miller and Wolford argued that the critical item elicits a very liberal criterion shift because it is highly related to the list items, resulting in a high false-alarm rate. They claimed that “a criterion shift is something that happens at retrieval, but is more of a decision process than a memory process per se” (p. 404). In other words, false remembering in this paradigm is due to erroneous decision processes (i.e., strategically guessing that critical items are *old*), and therefore is of limited relevance to the understanding of false memories in other contexts. Miller and Wolford did not explicitly state that these criterion shifts are strategic or within subjects’ control, although this is implicit in the argument. Indeed, if the proposed criterion shifts happen automatically at test, or are beyond subjects’ strategic control, it is unclear why such biases would not be considered an inherent part of the retrieval process (cf. Jacoby, Marsh, & Dolan, in press).

To support their argument, Miller and Wolford (1999) computed signal detection parameters and showed that a bias parameter was different for critical items than for list items or for unrelated items. However, both Wixted and Stretch (2000) and Wickens and Hirshman (2000) have pointed out that Miller and Wolford’s interpretation of the bias parameter as an index of a criterion shift represents a logical error. Within signal detection theory, changes in bias estimates can be due to either a shift in response criterion or to a shift in the underlying distributions. Thus, changes in bias estimates are not sufficient evidence for changes in response criteria, and no data exist in Miller and Wolford’s results to support one interpretation over the other.

We are grateful to Michael Miller and John Wixted for comments on an earlier draft and to James Neely and Patrick Dolan for correspondence on the subject. We also thank Hila Snyder for her assistance in collecting data. This research was supported by NIMH Grant 1RO3 MH59034-01 to K.B.M. and by a contract from the Office of Technical Services of the U.S. Government awarded to H.L.R. and K.B.M. Correspondence concerning this article should be addressed to D. A. Gallo, Department of Psychology, Campus Box 1125, Washington University, One Brookings Drive, St. Louis, MO 63130 (e-mail: dgallo@artsci.wustl.edu).

In what follows, we review the strongest evidence against this criterion shift account of DRM false memories and provide new evidence that effectively rules it out as a major explanation of the effect. In the present context, we use the term *criterion shift* to refer to intentional changes in a subject's inclination to say "old" to a test item. According to the theory in question, these changes are made on an item-by-item basis, on the basis of each test item's judged relatedness to a study list's theme (Miller & Wolford, 1999). In signal detection terms, such a criterion shift would mean that the subject requires less memorial evidence (e.g., trace strength or familiarity) to say "old" to test items that are highly related to a list (e.g., critical items) relative to other test items (e.g., unrelated lures). Such a shift would also occur for list items on the test if they were judged to be highly related to the theme of the list. In an extreme form of the theory, subjects would not use any memorial evidence to say "old" to test items, but would instead base their recognition decisions solely on a test item's perceived relatedness to a list. In this sense, critical false alarms would be due solely to strategic guessing, on the basis of judged relatedness, rather than on the basis of subjective familiarity and/or false recollection of the critical items. In either form of the theory, these criterion shifts are strategic, or under conscious control: Subjects explicitly assume that test items related to a list are likely to have been presented, and, on the basis of this assumption, they are more likely to respond "old" to these items relative to others.

Roediger and McDermott (1999) noted, among other considerations, that two warning studies rendered this strategic criterion shift interpretation highly improbable. Both Gallo, Roberts, and Seamon (1997) and McDermott and Roediger (1998) forewarned subjects about the paradigm, gave sample lists, and then urged the subjects to avoid false remembering when they studied and were tested on new lists. Gallo et al. (1997) found large reductions in critical false alarms following the warning, but these false alarms were still much higher than the false alarms to control items (whose associative list had not been studied). Gallo et al. (1997) suggested that warned subjects reduced false remembering by determining (i.e., "figuring out") some critical items at study and encoding them as "not presented." This claim was bolstered by (1) post-experiment questionnaires in which warned subjects reported using such strategies and (2) the finding that simply instructing uninformed subjects to be very cautious at test only resulted in a small, nonsignificant reduction in false recognition. Thus, merely encouraging subjects to be cautious was ineffective, which suggests that warned subjects had used more complex strategies to reduce false recognition.

In fact, the false recognition effect might have been underestimated in Gallo et al.'s (1997) warning condition because subjects were told that critical items were never studied. Armed with this knowledge, warned subjects could have reduced critical false alarms by responding "new" to any test item that they thought might be the critical item (i.e., the item most closely related to the list

items), regardless of whether they actually remembered this item from study.¹ McDermott and Roediger (1998) avoided the latter possibility. Their subjects were also given strong warnings about the illusion, but because the critical item appeared in half the lists, they could not increase accuracy by simply responding "new" to any item that might be the critical item. Instead, subjects had to carefully monitor their memories to determine whether the item had been presented. McDermott and Roediger did find some reduction in false recognition, but the reduction was not as great as that of Gallo et al. (1997). They also concluded that the illusion was remarkably robust and not easily overcome, much as is the case for perceptual illusions that persist even when people know that what they are seeing is not accurate.

We propose that it is unlikely (and logically inconsistent) that the high levels of false remembering in these fully informed subjects were caused by strategic criterion shifts. Informed subjects knew that test items that were highly related to a list's theme (such as the critical items) were not necessarily studied, and would therefore not have adopted liberal response criteria for such items. It follows that any residual false recognition effect in warning conditions must have been due to processes other than such criterion shifts. Of course, because the warnings of Gallo et al. (1997) and McDermott and Roediger (1998) did somewhat reduce false recognition, one might argue that part of the effect is caused by criterion shifts. However, this conclusion does not necessarily follow from these experiments. Because the warnings always occurred before study, the warning effects could have been due to encoding processes (as suggested by Gallo et al., 1997) or decision processes at test (criterion shifts), or to a combination of both.

The critical experimental condition required to examine strategic criterion shifts directly is to give a strong warning after study but before test. If subjects are informed about the paradigm, and if they are instructed to respond carefully to items that may be critical items, false recognition should be reduced to the extent that it is caused by liberal criterion shifts to these items. If the effect arises from other sources, such as a strong gist trace (Reyna & Brainerd, 1995) or associative activation (Roediger, Balota, & Watson, 2001; Underwood, 1965), the warning after study might not affect false recognition because by then the information leading to false recognition will have already been encoded. The present experiment was designed to provide a direct test of this strategic criterion shift account and also to compare directly the effects of strong warnings given either before or after the study phase.

Rather than attempting to infer criterion shifts from ambiguous signal detection estimates of bias, we took a more direct approach by experimentally manipulating strategies. We employed four conditions. In the three primary conditions, subjects heard 12 DRM lists, half of which contained the critical item. One was a standard condition in which subjects heard the lists and received a recognition test with instructions simply to be as accurate as possible (the no warning condition). On the basis of previous

research, we expected high levels of false recognition, approximating the hit rate, in this condition. A second condition provided a warning about the paradigm before study (the warning before study condition), as in Gallo et al. (1997) and McDermott and Roediger (1998). On the basis of these studies, we expected that the false recognition effect would be reduced (relative to no warning) but not eliminated in this condition.

A third condition, which was designed to test the strategic criterion shift account, provided subjects with a warning after study but before test (the warning after study condition). If liberal criterion shifts play a role in the false recognition effect under standard conditions, warned subjects should demonstrate reduced false recognition (relative to no warning) because they would not have adopted this criterion shift strategy. In fact, because warned subjects were told that the critical item was sometimes studied and other times it was not, they would have no incentive to adopt any guessing strategy for these items and instead should have based their recognition decisions solely on retrieved information. Conversely, if such criterion shifts play no role under ordinary circumstances, simply discouraging subjects from using a criterion shift strategy should not influence the false recognition effect. False recognition should be equal to that in the no warning condition.

Of course, if the warning after study did not affect false recognition, a critic might argue that the warning was simply ineffective at motivating subjects to abandon a liberal criterion shift strategy. To safeguard against this potential ambiguity, we included a fourth condition that was identical to the warning after study condition, except that the critical item was never presented in the study lists (the warning after study–no critical targets condition, hereafter warning after study–no CT). As was discussed previously, any reduction of false recognition in this condition would tell us little about the mechanisms driving false recognition. Because subjects were told that the critical items were never presented at study, they could reduce overt false recognition by calling any highly related item “new,” even if they were sure that it had been studied (i.e., they falsely remembered it). Nevertheless, this condition does serve as an important manipulation check. Telling subjects that highly related critical items were *never* presented and that they should say “new” to those items should motivate them to use *conservative* criterion shifts on any test item that they perceive as highly related to a list (i.e., one that might be the critical item). As a result, reduced false recognition in this condition would provide converging evidence that warned subjects were willing and able to adopt the response strategy that was appropriate for their instructional condition.

METHOD

Subjects

Ninety-six Washington University undergraduates participated for course credit, and 24 were arbitrarily assigned to each of the four conditions. Thirteen other subjects were replaced because they

indicated on a postexperiment questionnaire that they had previous knowledge of the paradigm.

Materials and Design

The 18 DRM lists that elicited the highest levels of false recall in the Stadler, Roediger, and McDermott (1999) norms were used, with the exception of substituting the *needle* list for the *sleep* list, which was given as an example to warned subjects. Each list consisted of the 15 most common words elicited by a stimulus word (i.e., the critical item) on a free association task. Study lists were digitally recorded in a female voice and presented over computer headphones at a rate of one word every 1.5 sec. Lists were preceded by a 5-sec visual prompt, indicating the number of the next list (e.g., List 1).

The 18 study lists were divided into three sets that were equated on the mean probability of false recognition (Stadler et al., 1999). Each subject studied 12 lists. In the three primary conditions, the critical item was studied in 6 lists, replacing the item from position 5 (3 lists) or position 9 (3 lists). Each list set was studied with the critical item, studied without the critical item, and not studied an equal number of times across subjects in each condition. Lists were presented in a random order with the exception that no 4 lists of the same type (critical item studied in position 3, position 9, or not studied) were presented in sequence. The warning after study–no CT condition was identical to the warning after study condition, except that the critical item was never studied in any of the lists and subjects were told this in the warning.

All subjects received the same 84-item visual recognition test. The critical item and the list items from positions 3, 7, and 11 were tested from each of the 18 lists. To equate the number of studied and nonstudied test items in the 3 primary conditions, the test also included the critical item and three list items from each of 3 lists in Stadler et al. (1999) that were not studied by any subject. Because critical items were never studied in the warning after study–no CT condition, the addition of these filler items resulted in 36 test items that had been presented at study and 48 that had not. In all conditions, test items were ordered randomly, with the exception that words associated with the same list were separated by at least three items.

Procedure

One to 4 subjects were tested at a time. The instructions are described here, and the complete set that was read to the subjects can be obtained from www.iac.wustl.edu/~mclweb or from the first author.

Prior to study, all subjects were told to pay close attention because they would hear 12 word lists presented quickly and that their memory would be tested. In the no warning and the two warning after study conditions, no mention was made of the nature of the word lists or of memory illusions. However, in the warning before study condition, the subjects were told that each word list would contain items related to one “common word.” They were read the *sleep* list and told that the common word, like *sleep*, would be presented in only some of the lists. It was explained that people often falsely remember the common word as having been on the list. To avoid this error, they were to identify the common word for each list and to note whether or not it had been presented (this was the most common strategy reported by informed subjects in Gallo et al., 1997).

All subjects were given test instructions immediately after the study phase. They were told that words would be presented on the computer screen and would be preceded by a test number that corresponded to that on their response sheet. They were told that some of the words were *old* (studied) and some were *new* (not studied). They were to circle the appropriate response (*old* or *new*). In addition, for each *old* word, they were to make a remember/know judgment by circling R or K. A *remember* judgment was to be made when they could recollect a specific detail about the word’s occurrence in the study list. A *know* judgment was to be made when they recognized the word but could not recollect any details about its prior occurrence. We did not include a *guess* option (e.g., Gardiner & Conway, 1999) because

the warning instructions discouraged subjects from guessing *old*. Each test item remained on the screen until the subject pressed the "enter" key.

The warning after study group was given a warning comparable to the warning before study. That is, the illusion was explained to subjects, they heard the *sleep* list as an example, and they were told that the common word had been presented in only some of the study lists. The warning after study–no CT group received similar instructions, except that they were told that the common word had never been presented. In all three warning conditions, test instructions stressed that a test item that fit well within the theme of the list *might* be one of the nonpresented common items and that the subjects should avoid responding "old" to these items. Thus, the subjects had no reason to use a liberal criterion shift strategy to critical items. The time to deliver test instructions was roughly equivalent across the four conditions (2–3 min), and the recognition test immediately followed.

RESULTS

Effects reported as significant reached the .05 level of confidence. Recognition data are presented in Table 1. The top four rows show data for studied lists in which the critical item appeared (critical item in list) or did not appear (critical item not in list). The next two rows show false alarms to items from these same lists when the corresponding lists were not presented (nonstudied-list controls). The last two rows show false alarms to filler items from lists that were never studied by any subject (nonstudied-list fillers). The data presented are mean hit or false-alarm rates and how these are decomposed into *remember* and *know* judgments (in parenthesis).

For simplicity, we first discuss the results from the three primary conditions and then discuss the warning after study–no CT condition. The key comparison is between critical false alarms when the list was studied (Table 1, Row 4) and false alarms to these items when the list was not studied (i.e., critical item controls, Row 6). The difference between these two represents the proportion of the false recognition effect that is due solely to the presentation of associates. In all three conditions, false alarms to

critical items were much higher when their study list was presented than when it was not [.88 vs. .31 with no warning, .61 vs. .33 with warning before study, and .79 vs. .26 with warning after study; smallest $t(23) = 4.01$, $SEM = 0.069$]. The 52% effect in the warning after study condition (i.e., .79–.26) was not reliably different from the 57% effect found in the no warning condition (.88–.31) [$t(46) < 1$], indicating that the warning provided after study had little effect. Because this warning precluded a liberal criterion shift strategy, such criterion shifts cannot account for this 52% false recognition effect.

We now present more complete statistical analyses of the data to bolster this main point and others. To facilitate comparisons across item types and conditions, we corrected recognition for list items and critical items for each subject by using several procedures (Snodgrass & Corwin, 1988). With the high-threshold correction procedure (as was used above for critical items), recognition of list items and critical items was adjusted by simply subtracting false alarms to list item controls and critical item controls, respectively. This correction provides an index of the influence of list presentation on recognition performance, factoring out the portion of the total probability of recognition that was due to unknown factors (e.g., a priori item differences and/or noise between experimental conditions, both of which are measured by false alarms to control items whose list was not presented). We also calculated two alternative corrections, one parametric (d') and one nonparametric (A' ; Donaldson, 1992).² In all corrected analyses reported hereafter, we report only the high-threshold corrected data, because all three corrections yielded the same pattern of results and conclusions reached by inferential statistics.

When the critical item was presented in the study list, there were no differences in corrected recognition of list items or critical items across the three primary conditions. Corrected recognition of list items was .55 (no warning condition), .57 (before study condition), and .53 (after

Table 1
Average Proportion of List Items and Critical Items Recognized (OLD) Along With Corresponding Remember (R) and Know (K) Responses for Each Warning Condition.

	Time of Warning							
	No Warning		Before Study		After Study		After Study–No CT	
	OLD	(R/K)	OLD	(R/K)	OLD	(R/K)	OLD	(R/K)
Critical item in list								
List items	.76	(.51/.25)	.81	(.53/.28)	.71	(.42/.29)	n.a.	n.a.
Critical items	.89	(.72/.17)	.90	(.67/.22)	.86	(.62/.24)	n.a.	n.a.
Critical item not in list								
List items	.70	(.48/.22)	.80	(.50/.30)	.71	(.45/.26)	.70	(.42/.28)
Critical items	.88	(.69/.19)	.61	(.40/.21)	.79	(.49/.30)	.66	(.38/.29)
Nonstudied-list controls								
List items	.21	(.07/.14)	.24	(.06/.18)	.18	(.03/.15)	.31	(.07/.24)
Critical items	.31	(.10/.21)	.33	(.09/.24)	.26	(.04/.23)	.36	(.11/.25)
Nonstudied-list fillers								
List items	.15	(.03/.13)	.18	(.07/.11)	.12	(.01/.12)	.16	(.03/.13)
Critical items	.22	(.04/.18)	.19	(.04/.15)	.17	(.04/.13)	.26	(.10/.17)

Note—n.a. = not applicable. Instances where remember and know judgments do not sum to the proportion of recognized items reflect rounding error.

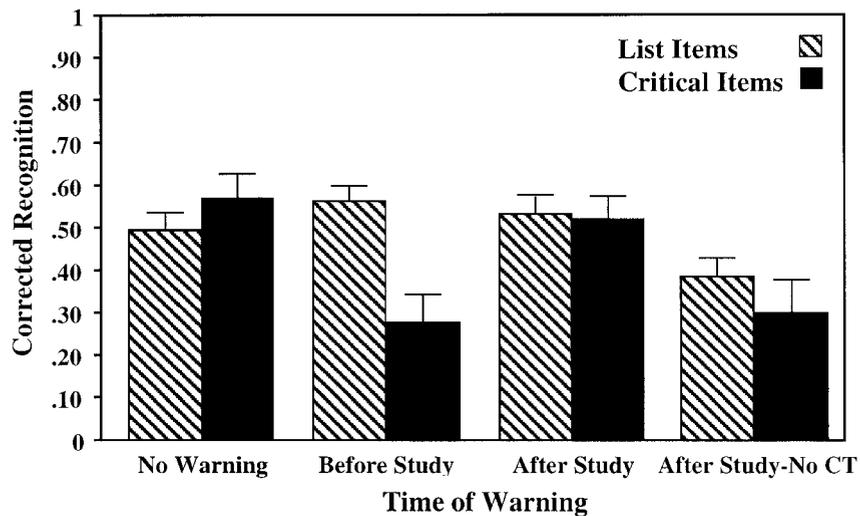


Figure 1. High-threshold corrected recognition (see text) of list items (hatched) and critical items (black) in each of the four instructional conditions. These data are from those lists in which the critical item was not presented during study. Error bars represent the standard error of each mean.

study condition). Corrected recognition of presented critical items was .58, .56, and .60, respectively. A 2 (item type) \times 3 (instruction) analysis of variance (ANOVA) indicated no effect of item type [$F(1,69) = 1.11$, $MS_e = 0.032$, $p > .20$], instruction [$F(2,69) < 1$], or interaction [$F(2,69) < 1$]. These results suggest that these warned subjects did not guess *new* to any item related to a list (i.e., did not make conservative criterion shifts, see Note 1), regardless of whether they actually remembered the item. Apparently, including the critical item in some lists successfully discouraged the subjects from relying on such a strategy.

When the critical item was *not* studied in the list, there were differences in corrected recognition across the three primary conditions (as is shown in Figure 1). A 2 (item type) \times 3 (instruction) ANOVA revealed a main effect of item type [$F(1,69) = 4.58$, $MS_e = 0.04$] and an interaction between item type and instruction [$F(2,69) = 10.41$, $MS_e = 0.04$]. There was no main effect of instruction [$F(2,69) = 2.56$, $MS_e = 0.075$, $p > .05$]. To explore the interaction, we compared corrected hits and critical false alarms for each warning condition to those of the no warning condition. For veridical recognition, there were no differences between the no warning condition (.49) and either the warning before study condition (.56) [$t(46) = 1.29$, $SEM = 0.052$, $p > .20$] or the warning after study condition (.53) [$t(46) < 1$]. As was the case when the critical item was presented in the study list, recognition of list items was relatively unaffected by warnings. In contrast, false recognition was affected by instruction: The warning before study considerably lowered false recognition (.28) relative to the no warning condition (.57) [$t(46) = 3.32$, $SEM = 0.088$], whereas the warning after study did not significantly reduce false recognition (.52)

relative to the no warning condition [$t(46) < 1$], as was previously noted.

Although the 5% effect of the warning after study was not significant, it was in the direction that one would predict if these subjects were more cautious to nonpresented critical items than were the subjects in the no warning condition. A similar pattern was observed in *remember* judgments. When these judgments are corrected for false alarms to control items in the same way as *old/new* judgments were, it can be seen that both the warning before study (.31) and after study (.45) reduced *remember* responses relative to the no warning condition (.59). This reduction was significant for the warning before study condition [$t(46) = 3.01$, $SEM = 0.089$], but not for the warning after study condition [$t(46) = 1.64$, $SEM = 0.085$, $p > .10$]. Thus, explicitly discouraging the warning after study subjects from using liberal criterion shifts had some effects, but these were only slight. More importantly, *remember* judgments were given more frequently to non-presented critical items than to critical-item controls in all three conditions [.69 vs. .10 with no warning, .40 vs. .09 with warning before study, and .49 vs. .04 with warning after study; smallest $t(23) = 4.91$, $SEM = 0.064$]. This again demonstrates robust false recognition in all three conditions.

One final issue of interest is the discriminability between critical items when they were presented in the list and when they were not. In the no warning condition, these probabilities did not differ either for overall recognition [.58 vs. .57; $t(23) < 1$] or for *remember* responses [.72 vs. .69, $t(23) < 1$]. Caution should be taken in interpreting this null result of presentation, though, since McDermott (1997), McDermott and Roediger (1998), and Miller and Wolford (1999) have all reported greater recall

or recognition of critical items when they are presented in lists than when they are not. The failure to find the effect here may be due to a ceiling effect in the uncorrected veridical and false recognition (.89 and .88), and inspection of individual subject's performance confirmed that 10 of the 24 subjects had both critical hit and false-alarm rates of 100%, or at ceiling. List item recognition was slightly greater when the critical item was in the list than when it was not [.55 vs. .49; $t(23) = 2.09$, $SEM = 0.027$], but this pattern did not extend to *remember* responses [.51 vs. .48; $t(23) < 1$].

Unlike the no warning condition, the subjects in the warning before study condition and the warning after study condition were able to distinguish between instances when the critical item was presented and not presented. A 2 (item type) \times 2 (instruction) ANOVA on recognition of the critical items in these two warning conditions indicated a main effect of item type [$F(1,46) = 41.19$, $MS_e = 0.019$] and an interaction between item type and instruction [$F(1,46) = 13.71$, $MS_e = 0.019$]. The main effect of instruction was not significant [$F(1,46) = 3.49$, $MS_e = 0.133$, $p > .05$]. The interaction indicates that the ability to discriminate between presented and non-presented critical items was greater with the warning before study [.56 vs. .28; $t(23) = 6.27$, $SEM = 0.045$] than after study [.60 vs. .52; $t(23) = 2.30$, $SEM = 0.033$], although the effect was significant in each case. This pattern was also evident in *remember* judgments. Finally, recognition of list items and remember judgments to list items did not depend on whether or not the critical item was in the list for either of the two warning conditions (all $t_s < 1$). Overall, these data indicate that the warning before study was much more effective at reducing false recognition than when the warning was withheld until test, when the effect was much smaller.

A critic might argue that the large false recognition effect in the primary warning after study condition simply indicates that this warning was ineffective at discouraging subjects from making liberal criterion shifts to critical items. However, several pieces of evidence suggest that these subjects were trying to avoid false recognition, in which case they would not have used such liberal criterion shifts. First, the warning after study condition did yield small but nonsignificant reductions in critical false alarms and *remember* judgments compared with the no warning condition. In addition, the subjects in the warning after study condition were able to discriminate between critical items that were in the list and those that were not in the list (using both *old/new* and *remember* judgments), whereas subjects in the no warning condition were not. Both of these findings suggest that subjects in the warning after study condition were influenced by the instructions. Second, and more importantly, there was a large effect of warning when it was given before study. This demonstrates that warned subjects certainly were willing to reduce false recognition. Thus, the failure to find a large reduction in the warning after study condition suggests that these subjects were unable (as opposed to unwilling) to reduce false recognition considerably.

We turn lastly to whether our giving subjects incentive to use conservative criterion shifts to critical items (as opposed to their avoiding criterion shifts altogether) can reduce the false recognition effect. The warning after study–no CT condition was designed to explore this possibility. Because these subjects were told that the critical item was never studied, they could have reduced overt false recognition by simply responding conservatively to any item that might have been a critical item. Note that this manipulation would selectively reduce critical false alarms only to the degree that the subjects could identify critical test items as such. If subjects also judged some list items to be highly related to a list's theme, they would respond more conservatively to these items, too. As discussed previously, reductions in false recognition in this condition would be ambiguous with respect to the mechanisms that cause false recognition, but they would provide an important manipulation check for our primary warning after study condition.

The results from the warning after study–no CT condition can be found in the fourth column of Table 1 and were consistent with our expectation. Uncorrected false recognition of critical items was marginally lower in this condition (.66) than in the original warning after study condition (.79), [$t(46) = 1.86$, $SEM = 0.067$, $p = .07$], and the same pattern was found with *remember* responses (.38 vs. .49). We again corrected the *old/new* data by subtracting false alarms to control items, and the resulting data can be found in Figure 1. As can be seen, corrected false recognition was significantly lower in the warning after study–no CT condition relative to the original warning after study condition [.30 vs. .52; $t(46) = 2.39$, $SEM = 0.093$], as was veridical recognition [.38 vs. .53; $t(46) = 2.69$, $SEM = 0.055$]. These data suggest that the subjects were biased to say "new" to any items that appeared to be related to the list (i.e., to critical items and, to a lesser extent, list items) and provide additional evidence that our warning after study instructions did influence response strategies. Finally, in the warning after study–no CT condition, uncorrected false recognition of critical items (.66) was significantly greater than that for nonstudied control items (.36) [$t(23) = 3.88$, $SEM = 0.077$], and a similar effect was found for *remember* judgments [.38 vs. .11; $t(23) = 5.44$, $SEM = 0.049$]. Thus, even though this condition could potentially underestimate the false recognition effect, false recognition persisted at high levels.

DISCUSSION

The primary conclusion from our results is that strategically liberal criterion shifts made specifically to critical items at test play little or no role in the DRM false recognition effect. In the no warning condition, corrected false recognition of critical items (.57) was about the same as corrected recognition of list items (.49), or even slightly higher, demonstrating the standard effect. When the warning was given before study, the subjects were able to reduce false recognition relative to veridical recognition (.28 vs. .56), replicating past research (e.g., Gallo et al., 1997;

McDermott & Roediger, 1998). However, in the warning after study condition, the subjects were not able to reduce false recognition of critical items relative to list items (.52 vs. .53). With the warning before study, we assume that, for at least some lists, the subjects identified the critical item and the fact that it was not studied and later used this information to reduce false recognition. This strategy could not be used when the warning was given after study, and these subjects demonstrated very high levels of false recognition even though they were not using liberal response criteria for these critical items. Clearly, if such liberal criterion shifts play any role at all in the DRM false recognition phenomenon, it is a very small one.

Although our primary warning after study was mostly ineffective, we do not mean to imply that all warnings given after study would be ineffective at reducing false remembering. The efficacy of a warning will depend on the type of strategies that it encourages subjects to use and on the types of information in memory that will help the subject discriminate between what was studied and what was not. For instance, the warning after study–no CT condition demonstrated that conservative criterion shifts can somewhat reduce critical false alarms, although the subjects apparently made such shifts only because they were assured that highly related items were never studied. (In the typical DRM procedure, the subject is not told this information). As was discussed, the effects of this warning might only indicate that the subjects circled *new* to some highly related test items (including critical and list items), regardless of their memory for these items, in an effort to follow instructions and not make false alarms. Anastasi, Rhodes, and Burns (2000, Experiment 3) recently provided results that are relevant to this issue. These authors used a warning comparable to that in the warning after study–no CT condition and found that warned subjects were less likely than naive subjects to make *remember* judgments to both critical items (.36 vs. .59) and list items (.55 vs. .62), although statistical analyses of these comparisons were not reported. Our results are consistent with these, in that the warning after study–no CT condition demonstrated reduced recognition of both critical and list items. Taken together, perhaps the most surprising finding from these conditions is that the false recognition effect remained largely intact, especially in light of the fact that the effect might be underestimated with warnings under these conditions.

Other warnings given after study might be even more effective at reducing false recognition if they provide strategies that help subjects monitor the source of the information that can cause illusory memories (e.g., Hicks & Marsh, 1999). However Neuschatz, Payne, Lampinen, and Toglia (2001) have failed to find an effect even with these types of warnings. This is not to say that subjects cannot use monitoring strategies at test to reduce false remembering. Even with naive subjects, we (e.g., Gallo, McDermott, Percer, & Roediger, 2001; McDermott & Watson, 2001) and others (e.g., Benjamin, 2001; Schacter, Israel, & Racine, 1999) have demonstrated that en-

hancing subjects' memory for list items sometimes helps them reduce false remembering, apparently via monitoring strategies. The main point to be made with our primary warning after study is that instructing subjects to avoid the illusion, thereby precluding liberal criterion shifts for critical items, is not sufficient to reduce false recognition.

In sum, the evidence indicates that the DRM false recognition effect is not caused by decisions to make liberal criterion shifts to critical items at test. Other candidate theories include fuzzy trace theory (Brainerd & Reyna, in press; Reyna & Brainerd, 1995) and activation/monitoring theory (Roediger, Balota, & Watson, 2001; Roediger & McDermott, 1995; Roediger, Watson, McDermott, & Gallo, 2001). The former postulates that gist representations and specific, or verbatim, representations combine to determine performance, whereas the latter hypothesizes that activation of the critical item and subsequent monitoring of this activation are two critical sets of concerns. The present results do not test between these two theories, but they do constrain each of them. Specifically, the robust false recognition effect in the warning after study condition indicates that the influence of information that causes false recognition (e.g., a gist representation or an activated representation of the critical item itself) is beyond subjects' conscious control. As a result, in the absence of additional memorial information and/or strategies to facilitate monitoring, it may be impossible for subjects to avoid false memories by merely willing them away. This conclusion may sound discouraging, but it makes sense if one considers that many false memories are caused by inherent mechanisms that, under most circumstances, allow us to make useful inferences and associations and to extract meaningful information from our environments. The important implication of our results is that the minimization of such false memories is best achieved by controlling both what is learned at encoding and how this information is used at retrieval, rather than trying to avoid false remembering solely at the time of retrieval.

REFERENCES

- ANASTASI, J. S., RHODES, M. G., & BURNS, M. C. (2000). Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. *American Journal of Psychology*, **113**, 1-26.
- BENJAMIN, A. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 941-947.
- BRAINERD, C. J., & REYNA, V. F. (in press). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development & Behavior*.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.
- DONALDSON, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, **121**, 275-277.
- GALLO, D. A., MCDERMOTT, K. B., PERCER, J. M., & ROEDIGER, H. L., III (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 339-353.

- GALLO, D. A., ROBERTS, M. J., & SEAMON, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, **4**, 271-276.
- GARDINER, J. M., & CONWAY, M. A. (1999). Levels of awareness and varieties of experience. In B. H. Challis & B. M. Velichkovsky (Eds.), *Stratification in cognition and consciousness* (pp. 237-254). Amsterdam: John Benjamin.
- 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.
- JACOBY, L. L., MARSH, E. J., & DOLAN, P. O. (in press). Forms of bias: Age-related differences in memory. In M. Naveh-Benjamin, M. Moscovitch, & H. L. Roediger III (Eds.), *Perspectives on human memory and cognitive aging: Essays in honour of Fergus Craik*. Hove, U.K.: Psychology Press.
- McDERMOTT, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory & Language*, **35**, 212-230.
- McDERMOTT, K. B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin & Review*, **4**, 582-586.
- McDERMOTT, K. B., & ROEDIGER, H. L., III. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*, **39**, 508-520.
- McDERMOTT, K. B., & WATSON, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory & Language*, **45**, 160-176.
- MILLER, M. B., & WOLFORD, G. L. (1999). Theoretical commentary: The role of criterion shift in false memory. *Psychological Review*, **106**, 398-405.
- NEUSCHATZ, J. S., PAYNE, D. G., LAMPINEN, J. M., & TOGLIA, M. P. (2001). Assessing the effectiveness of warnings and phenomenological characteristics of false memories. *Memory*, **9**, 53-71.
- 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.
- REYNA, V. F., & BRAINERD, C. J. (1995). Fuzzy trace theory: An interim synthesis. *Learning & Individual Differences*, **7**, 1-75.
- ROEDIGER, H. L., III, BALOTA, D. A., & WATSON, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Suprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington, DC: American Psychological Association.
- 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.
- ROEDIGER, H. L., III, McDERMOTT, K. B., & ROBINSON, K. J. (1998). The role of associative processes in creating false memories. In M. A. Conway, S. E. Gathercole, & C. Cornoldi (Eds.), *Theories of memory II* (pp. 187-245). Hove, U.K.: Psychological Press.
- ROEDIGER, H. L., III, WATSON, J. M., McDERMOTT, K. B., & GALLO, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychological Bulletin & Review*, **8**, 385-407.
- SCHACTER, D. L., ISRAEL, L., & RACINE, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory & Language*, **40**, 1-24.
- SNODGRASS, J. G., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.
- STADLER, M. A., ROEDIGER, H. L., III, & McDERMOTT, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, **27**, 494-500.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychologist*, **26**, 1-12.
- UNDERWOOD, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, **70**, 122-129.
- WICKENS, T. D., & HIRSHMAN, E. (2000). False memories and statistical decision theory: Comment on Miller and Wolford (1999) and Roediger and McDermott (1999). *Psychological Review*, **107**, 377-383.
- WIXTED, J. T., & STRETCH, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review*, **107**, 368-376.

NOTES

- Note that this is a similar sort of criterion shift to test items that Miller and Wolford (1999) proposed. The key difference is that they proposed that subjects spontaneously shift criteria under standard conditions, and the shift is in the liberal direction (thereby causing false alarms). The criterion shift discussed here is allegedly caused by warnings, and is in the conservative direction (thereby reducing false alarms that would otherwise be caused by memorial processes).
- To avoid hit rates of 0 or 1 these estimates were based on transformed recognition data, $p(\text{hit or FA}) = (x + .5)/(n + 1)$. The interested reader may find these estimates posted on the same Web page as our instructions.

(Manuscript received March 21, 2000;
revision accepted for publication November 22, 2000.)