MEMORY, 2006, 14 (1), 115–126

Psychology Press
Taylor & Francis Group

Re-exposure to studied items at test does not influence false recognition

Michael D. Dodd, Erin D. Sheard, and Colin M. MacLeod

University of Toronto, Canada

In two experiments, we investigated whether re-exposure to previously studied items at test affects false recognition in the DRM paradigm. Furthermore, we examined whether exposure to the critical lure at test influences memory for subsequently presented study items. In Experiment 1, immediately following each studied DRM list, participants were given a recognition test. The tests were constructed such that the number of studied items preceding the critical lure varied from zero to five. Neither false recognition for critical lures nor accurate memory for studied items was affected by this manipulation. In Experiment 2, we replicated this pattern of results under speeded conditions at test. Both experiments confirm that exposure to previously studied items at test does not affect true or false recognition in the DRM paradigm. This pattern strongly suggests that retrieval processes do not influence false recognition in the DRM paradigm.

Over the past decade, numerous studies have been conducted using a paradigm for generating false memories created by Deese (1959) and expanded by Roediger and McDermott (1995)—now known as the Deese-Roediger-McDermott (DRM) paradigm (e.g., Benjamin, 2001; Dodd & MacLeod, 2004; Roediger & McDermott, 1995; Seamon, Luo, & Gallo, 1998). The procedure underlying this paradigm is quite simple: Participants are presented with a series of lists to learn for a subsequent memory test, normally recall and/or recognition. For each list, all of the presented words converge on one unpresented associate (the critical lure: e.g., the list items *hot*, *winter*, and *ice* are all related to *cold*). Generally, this procedure leads to high rates of both false recall and false recognition of the critical lure (Deese, 1959, Roediger & McDermott, 1995).

Two theories have attempted to account for the source of these false memories—fuzzy trace (Brainerd & Reyna, 1996, 1998) and activation/monitoring (Roediger, Watson, McDermott, & Gallo, 2001), both proposing that the creation of false memories is attributable to processes occurring at encoding. However, in their seminal article Roediger and McDermott (1995) posited that retrieval processes might also play a role in the creation of false memories. Specifically, they noted that the false recall of a critical lure often occurred after many actually studied items had already been recalled. Moreover, at recognition, the critical lure was often preceded by the presentation of many studied items. This feature of testing left open the possibility that activation of studied items at retrieval combined with activation of these same items at encoding in creating a false memory for the critical lure. Although this test-based account is not inconsistent with activation-based accounts of false memory, which theorise that false memories can be created via a spread of activation to semantic associates when related words are studied (e.g., activation monitoring), it has received scant attention in the literature. Recently, however, Marsh, McDermott,

Correspondence should be addressed to Michael D. Dodd, Department of Psychology, 100 St. George Street, University of Toronto, Toronto, Ontario, M5S 3G3, Canada. Colin M. MacLeod is now at the Department of Psychology, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada. Email: michaeld@psych.utoronto.ca

This research was supported by discovery grant A7459 from the Natural Sciences and Engineering Research Council of Canada and by NSERC postgraduate scholarships. We thank Karen Roberts for her assistance in collecting the data. We also thank John Seamon and an anonymous reviewer for many helpful comments on an earlier version of this manuscript.

http://www.psypress.com/memory
DOI:10.1080/09658210444000575

and Roediger (2004) reported three experiments suggesting that test-induced priming may play no role in the creation of false memories.

Marsh et al. (2004) tested the influence of test-induced priming by manipulating the number of related items that preceded the critical lure at test. In their Experiment 1, which is most relevant to the present research, participants studied 18 15-item DRM lists, with the items presented visually on a computer monitor one by one and read aloud by the experimenter. Participants were instructed to memorise as many of the presented items as they could for a later memory test. Following the study session, participants completed a recognition test in which 297 words were presented to them, one at a time, and they were to decide whether each word had appeared in an earlier studied list by circling ''old'' or ''new'' on an accompanying response sheet. Remember/know judgements (Roediger & McDermott, 1995; Tulving, 1985) were required for ''old'' responses. The recognition test was constructed such that the critical lure followed the presentation of either zero, three, or six of the previously studied items, the expectation being that individuals might be more likely to endorse a critical lure when it had been preceded by three or six previously studied items at test than when that critical lure was the first item tested.

There was no evidence of test-induced priming at recognition: In particular, it did not matter how many actually studied list words preceded the critical lure from each list. Nor was there any difference in remember/know responses as a function of the number of studied items preceding the critical lure at test. Similar results were obtained in two subsequent experiments in which memory was tested via cued recall. Overall, then, the Marsh et al. (2004) study provides initial evidence that test-induced priming plays little role in the creation of false memories in the DRM paradigm.

The present research was designed to evaluate further the possible role of test-induced priming in false recognition. We sought to extend the study of test-induced priming in three key ways. First, we tested recognition following the presentation of each individual DRM list. In contrast, Marsh et al. (2004, Experiment 1) had presented one large recognition test following presentation of all 18 lists. It could well be that this long, delayed test was not sensitive to list-specific effects that would be apparent on an immediate test of each list. Second, we tested recognition under both the usual unspeeded test conditions (Experiment 1) and under speeded test conditions (Experiment 2). We reasoned that vulnerability to critical lures as a function of prior tested study items might be more likely when making decisions quickly, given that response time pressure has been shown to reduce source memory (e.g., Benjamin & Craik, 2001). Finally, we investigated one additional question—whether exposure to the critical lure at test influences correct memory for actually presented items. This reverses the test-priming question: If exposure to studied items at test could increase false memory via spreading activation to critical lures, then it seems reasonable that presenting the critical lure at test could also lead to a spreading of activation to actually studied semantic associates. This possibility can be investigated by examining correct recognition of studied items following the presentation of the critical lure.

EXPERIMENT 1

In Marsh et al.'s (2004) Experiment 1, studied items were presented visually and all 18 DRM lists were studied prior to a single ''batch'' test. In contrast, in the present experiment, items were presented auditorily, following more common practice with this procedure (e.g., Kellogg, 2001; Roediger & McDermott, 1995). Auditory presentation typically leads to higher rates of false recall/recognition relative to visual presentation, possibly because visual presentation provides distinct orthographic cues that aid in differentiating between actually presented items and critical lures (for which no orthographic information should exist despite the semantic association; Kellogg, 2001; Smith & Hunt, 1998). The absence of orthographic cues for items studied under auditory presentation might increase the likelihood of test-induced priming influencing false recognition. Reducing the ease of rejecting a critical lure could make a possible retrieval component of false recognition more apparent.

Furthermore, in the present experiment, recognition memory was tested following each individual DRM list instead of following all lists. There are two advantages to this procedure. First, in some studies using the DRM paradigm, memory has been tested following the presentation of each list (e.g., Roediger & McDermott, 1995; Stadler, Roediger, & McDermott, 1999). Consequently, the present experiment (in combination with those of Marsh et al., 2004) will provide a more complete understanding of the role of test-induced priming in the creation of false memories under varying test conditions. Although Marsh et al. (2004) did not observe test-induced priming for critical lures from studied lists, they did observe evidence of test-induced priming for critical lures from unstudied lists, which leads us to believe that there may be other conditions under which test-induced priming occurs. The present methodology allows us to test this possibility. Second, and more crucial, when memory is tested immediately, the activation levels for studied items should still be relatively high when the items are re-presented during the recognition test. In the Marsh et al. study, the long delay between study and test of individual items increased the probability that the activation level of actually studied items had been reduced substantially, and that re-exposure to these items might not increase activation beyond that observed at study.[1]

Hancock, Hicks, Marsh, and Ritschel (2003) have recently demonstrated that activation levels are key to the production of false memories. If the long lag between study and test in the Marsh et al. (2004) study led to a substantial decrease in the activation levels of both actually studied items and critical lures, then this delay may have neutralised any effect of re-exposure to these items at test. The removal of this delay between study and test in the present experiment allows us to investigate the effect of activation of the studied words and, perhaps, of the critical lure on the recognition test when activation levels are at their highest, thereby providing the most sensitive procedure possible.

The critical manipulation in the present experiment was the number of actually studied items that preceded the critical lure at test. If test-induced priming influences false recognition, then a difference in the level of false recognition would be expected as a function of the number of studied items preceding the critical lure at test. If activation levels are critical to the creation of false memories, then we would expect to observe test-induced priming with the present methodology, contrary to the results of Marsh et al. (2004). Moreover, the present methodology allowed us to examine not just test-induced priming as it relates to false memory, but also the effect of exposure to the critical lure on correct recognition. Specifically, we can determine the effect of exposure to the critical lure at test on memory for subsequently presented, actually studied items: Does false memory influence veridical memory?

Method

Participants. A total of 35 naïve students from the University of Toronto at Scarborough took part individually in a 30-minute session, receiving bonus points in Introductory Psychology for their participation.

Apparatus. Experimental programs, written in QuickBASIC 4.5, used the Graves and Bradley (1991) routines to achieve millisecond timing. An IBM-compatible 486 computer displayed the stimuli on a Magnavox 15″ colour monitor. All items were presented in standard DOS 80-character lowercase font in white on a black background. Participants made recognition decisions using two response keys (‘‘/’’ or ‘‘z’’, representing ‘‘old item’’ and ‘‘new item,’’ respectively).

Materials. A total of 12 of the DRM lists (*anger*, *chair*, *cold*, *cup*, *doctor*, *high*, *rough*, *smell*, *smoke*, *sweet*, *trash*, *window*) known to elicit high levels of false recognition were selected from the Stadler et al. (1999) norms. For each list, the 12 strongest associates of the critical lure were selected and recorded onto an auditory cassette tape in a male voice at a rate of one item every 2

[1] There is reason to believe that a long lag between the study of an item and a memory test for that same item would lead to a reduction in activation. For example, numerous studies of semantic priming have demonstrated a processing advantage for an item that follows a semantic associate (e.g., in a lexical decision task, participants are faster to identify the item ‘‘butter’’ as a word when it is preceded by a semantic associate such as ‘‘bread’’, relative to when it is preceded by an item such as ‘‘nurse’’). This priming is thought to be attributable to spreading activation to semantic associates when an item is processed. However, it has been demonstrated that semantic priming is ordinarily eliminated over long lags—e.g., when the prime (bread) and target (butter) are separated by 15 or more items—suggesting that activation decreases over time (McNamara, 1992; Ratcliff & McKoon, 1988). Moreover, MacLeod and Nelson (1976) had participants perform a continuous recognition task in which the lag between semantic associates was manipulated. Participants were very likely to false alarm to a semantic associate of an actually presented item at short lags (up to about five items) but false alarms reduced to baseline levels at long lags. Again, this is likely attributable to an initially high level of activation spreading to semantic associates but then decreasing over time. The present manipulation of testing following each list should address the possible decrease in activation that may have occurred in the Marsh et al. (2004) study when memory was tested following the presentation of a long series of lists.

seconds. Items within each list were always presented in the same order—from the strongest to the weakest associate—following common practice with this procedure (e.g., Dodd & MacLeod, 2004; Roediger & McDermott, 1995). The lists were presented to participants via a tape recorder. Presentation order of the 12 DRM lists was random.

A recognition test immediately followed the presentation of each list and consisted of 10 items: 5 actually presented items from the just-studied DRM list (items 1, 3, 5, 8, and 10); 2 weakly related/unstudied items (items 13 and 14 from the norms for that list); 2 unrelated/unstudied items from other DRM lists, and the critical lure from the just-studied list. The recognition tests were constructed such that the number of actually presented list items preceding the critical lure varied from zero to five. Of the 12 tests (one for each DRM list), there were two tests for each condition: 0, 1, 2, 3, 4, or 5 actually presented items preceding the critical lure. The presentation order of items at test was otherwise random with the only other restriction being that weakly related/unstudied items never preceded the critical lure. List presentation and test type were counterbalanced across participants.

Procedure. Participants sat about 30 cm from the monitor, responding via the keyboard. They were told that they would be hearing short lists of items auditorily and that at the end of each list they would be given a recognition test in which they were to decide as each item was presented on the screen whether that item had been on the study list. The recognition test was self-paced: Each item remained in the centre of the screen until the participant made a response. Upon response, the item was replaced by a row of white asterisks, indicating that the next item would appear in 500 ms. This procedure was repeated for each of the 12 lists, at the end of which participants were debriefed.

Results and discussion

The proportions of ''old'' responses, representing correct recognition of actually studied DRM list items and false recognition of weakly related/unstudied items, unrelated/unstudied items, and critical lures, are presented in Table 1. The proportion of ''old'' responses is reported as a function of the number of actually presented DRM list items (0–5: hereafter referred to as ''items preceding'') that preceded the critical lure at test. Corrected false recognition scores (which were calculated using a high threshold procedure: subtracting false alarms to unrelated/unstudied items from false alarms to critical lures) and standard errors

TABLE 1
Experiment 1

	Probability of an ''Old'' response					
	Items preceding					
Item type	*0*	*1*	*2*	*3*	*4*	*5*
Critical lures	.76	.81	.73	.86	.54	.71
	(.06)	(.05)	(.06)	(.05)	(.07)	(.06)
Presented items	.88	.85	.88	.83	.87	.88
	(.02)	(.02)	(.02)	(.03)	(.02)	(.03)
Weakly related/unstudied items	.17	.15	.09	.22	.16	.15
	(.04)	(.03)	(.03)	(.03)	(.03)	(.04)
Unrelated/Unstudied items	.06	.01	.01	.04	.01	.01
	(.02)	(.01)	(.01)	(.02)	(.01)	(.01)
Corrected critical lures	.70	.80	.72	.81	.53	.71
	(.06)	(.05)	(.06)	(.06)	(.07)	(.06)

Experiment 1: Mean probability of an ''Old'' response and Standard Error (SE) on the recognition test as a function of item type and of the number of studied items preceding the critical lure at test. Scores represent hits for presented items, and false alarms for critical lures, weakly related/unpresented items, and unrelated/unstudied items.

The final row reports corrected false recognition scores for critical lures, calculated as the false alarm rate to critical lures minus the false alarm rate to unrelated/unstudied items.

are also reported. For all subsequent analyses, corrected recognition scores are reported.[2]

To determine whether test-induced priming influences false recognition, mean proportions of ''old'' responses to critical lures were analysed using a one-way repeated measures analysis of variance (ANOVA). The effect of items preceding was significant, $F(5, 170) = 4.00$, $MSE = .09$, $p < .01$, although this was not attributable to any systematic difference in the likelihood of false alarms to critical lures as a function of items preceding. Rather, this was attributable to an unusually low number of false alarms to critical lures when four studied items preceded the critical lure. This anomaly is likely due to the low number of observations in each condition (only two critical lure observations per items preceding condition); we will address this peculiarity in Experiment 2. Indeed, when the condition with four previously studied items was removed from the analysis, there was no effect of items preceding on false recognition of critical lures, $F(4, 136) = 1.12$, $MSE = .09$, $p > .35$.

Association strength has also been shown to influence false recognition (e.g., Deese, 1959; Roediger et al., 2001). Therefore we investigated this factor in the present experiment for the cases where one associate preceded the critical lure and where two associates preceded the critical lure.[3] False recognition was unaffected by whether the critical lure was preceded by a weak associate(s) or a strong associate(s), $t(34) = 1.10$, $SE = .07$, $p > .25$, and $t(34) = 0.29$, $SE = .10$, $p > .75$, for the 1 and 2 associates preceding conditions, respectively.

To determine whether exposure to the critical lure influenced correct recognition of studied items, an analogous ANOVA was carried out on the mean proportions of ''old'' responses to actually studied items. There was no effect of items preceding, $F(5, 170) < 1$, indicating that correct recognition was not influenced by exposure to the critical lure at test. Remarkably, it did not matter whether the critical lure appeared prior to the presentation of all studied items (0 condition) or whether the critical lure appeared following the presentation of all studied items (5 condition): Correct memory in these two conditions was identical (.88). One might have imagined that the critical lure, being the word most associated to all of the actually studied words, would promote additional activation of studied words. That it did not, further emphasises the absence of test-based influences in the DRM paradigm.

Although not of primary interest here, analogous ANOVAs were also performed on the proportion of ''old'' responses to weakly related/unstudied items and to unrelated/unstudied items. There was a significant effect of items preceding the critical lure on both item types: $F(5, 170) = 2.22$, $MSE = .03$, $p < .05$, for weakly related/unstudied items; $F(5, 170) = 2.56$, $MSE = .03$, $p < .03$, for unrelated/unstudied items. As with critical lures, however, there was no systematic difference in ''old'' responses as a function of the number of studied items that preceded the critical lure.

The results of Experiment 1 are clearly in accord with those of Marsh et al. (2004)—differential test-induced priming did not influence false recognition of critical lures. The present research extends this finding to the situation where lists are presented auditorily (a manipulation known to increase false memory relative to visual presentation) and where the recognition test immediately follows each list (a manipulation expected to maintain the high activation level of studied

[2] Corresponding analyses were performed on uncorrected false recognition scores and on corrected false recognition scores, the latter determined by subtracting the false alarm rate for unrelated/unstudied items from the false alarm rate for critical lures. The pattern of results was identical, so for brevity we report only the analyses on corrected scores.

[3] For the association strength analyses, we only analysed trials on which one or two actually presented items preceded the critical lure. For these analyses, we treated the two strongest associates of the critical lure (items 1 and 3) as strong and the two weakest associates (items 8 and 10) as weak. Therefore, in the 1 item preceding condition, we analysed tests on which the critical lure was preceded by item 1 or 3, and compared this to tests on which the critical lure was preceded by item 8 or 10. For the 2 item preceding condition, we analysed tests on which items 1 and 3 preceded the critical lure, and compared this to tests where items 8 and 10 preceded the critical lure. This allowed for a pure comparison of association strength when one or two strong associates preceded the critical lure relative to one or two weak associates. When there are three or more studied items preceding the critical lure, it is difficult to determine the effect of association strength because any strong/weak comparison would contain the same studied items in both tests (e.g., 1, 3, and 5 preceding vs 5, 8, and 10 following). Furthermore, it would be difficult to determine whether the influence of association strength is cumulative (e.g., having just seen a strong associate, is there any additional effect of also seeing a weak associate prior to the presentation of the critical lure?), which creates further interpretation problems. Nonetheless, we are confident in concluding that association strength did not influence the present results given that there was no effect of association strength for the 1 and 2 items preceding conditions.

words during test). These changes should have increased sensitivity to test-based effects, yet no test-based effects were found. Moreover, we did not observe any effect on actually studied words of exposure to the critical lure at recognition: There was no difference in how individuals responded to actually presented study items that preceded, as opposed to followed, the critical lure. Taken together, the results of Experiment 1 suggest that the genesis of both true and false memories in the DRM paradigm is at encoding, and that the composition and sequence of the recognition test has no influence on memory performance in this situation.

EXPERIMENT 2

Although the results of Experiment 1 are consistent with the notion that test-induced priming does not influence false recognition, an alternate possibility needs to be addressed. It is possible that participants employed a response strategy that countered the effect of test-induced priming at recognition. For example, Brainerd, Reyna, Wright, and Mojardin (2003) have reported a phenomenon termed ‘‘recollection rejection’’ in which false alarms do not occur to critical lures that directly follow actually studied items at test. This may be due to the participant’s ability to reject a lure because it lacks the same contextual details that had been associated with the items seen earlier on the test (e.g., critical lures may be rejected because, unlike actually studied words, they do not elicit a verbatim trace). If this were the case, a process such as recollection rejection could counter any potential effects of test-induced priming.

The purpose of Experiment 2 was therefore twofold. First, we increased the number of observations per condition (from two to three) in an attempt to clarify the anomalous finding from Experiment 1 that false recognition of critical lures decreased significantly only when four studied items preceded the critical lure (but not when any other number of items preceded the critical lure). Second, and most importantly, to counter the possibility that recollection rejection (or a similar process, such as monitoring, see Marsh et al., 2004) influenced the results of Experiment 1, we attempted to create a testing situation in which participants had reduced access to an elaborative retrieval strategy. Specifically, we added response time pressure to the recognition test.

There is evidence that response time pressure might increase false memory. Benjamin (2001) has argued that recognition memory entails two processes. The first is a fast-acting spread of activation through semantic neighbourhoods to both actually studied items and semantic associates (e.g., the critical lure), evoking feelings of familiarity for these items. The second is more controlled and involves an attempt to determine the source of the familiarity and to set a decision criterion for the recognition response. In Benjamin’s study, young and older participants were presented with DRM list items at study—each list was presented either once or three times—and were tested under either speeded or unspeeded conditions. Intriguingly, false alarms to critical lures decreased with list repetition under unspeeded conditions, but increased under speeded conditions. Presumably, the speeded recognition decision forced participants to rely more on the automatic process of familiarity and less on the controlled processes. Because familiarity increased with list repetition, so too did false alarms to critical lures. Under unspeeded conditions, however, false alarms to critical lures actually decreased because participants were now able to make a controlled decision regarding the source of item familiarity.

If speeded recognition leads to reliance on the more automatic process of familiarity, then it is reasonable to posit that test-induced priming might influence false recognition of critical lures under speeded conditions. Specifically, increasing the number of studied items presented prior to the critical lure should increase the familiarity of the lure and, therefore, should lead to an increase in false recognition. Furthermore, if recollection rejection (or a similar process) influenced the pattern of results obtained in Experiment 1, response time pressure should eliminate the potential effects of this retrieval strategy, because participants will be forced to rely on activation (or gist) in responding to items at test. However, if test-induced priming plays no role in the creation of false memories, then we would expect to replicate the results of Experiment 1, with no difference in false recognition as a function of the number of actually studied items preceding the critical lure.

Method

Participants. A total of 30 naïve students from the University of Toronto at Scarborough took part individually in a 30-minute session, receiving bonus points in Introductory Psychology for their participation.

Apparatus. The apparatus was identical to that used in Experiment 1.

Materials. The materials were identical to those used in Experiment 1 except that six further DRM lists were added (*mountain*, *music*, *needle*, *slow*, *soft*, *thief*) to increase the number of observations per condition (number of items preceding the critical lure at test, 0–5) from two to three.

Procedure. The procedure was identical to that of Experiment 1 with the exception of additional response time pressure during the recognition test. Participants were informed that they would have only 750 ms to respond to each item as it appeared on the screen. They were explicitly instructed not to guess, but rather to decide whether the item was old or new the moment it appeared. If a response was made after 750 ms, an error tone was rapidly repeated 10 times, accompanied by the words ‘‘TOO SLOW’’ on the computer screen. However, in accord with Benjamin (2001), all responses were included in subsequent analyses, including those exceeding 750 ms. Participants had little difficulty responding within 750 ms and, on average, did not encounter the error tone more than twice per recognition test.

Results and discussion

The proportions of ‘‘old’’ responses, representing correct recognition of actually studied DRM list items and false recognition of weakly related/unstudied items, unrelated/unstudied items, and critical lures, are presented in Table 2. Corrected false recognition scores and standard errors are also reported and corrected false recognition scores were used for all subsequent analyses. As in Experiment 1, the proportions of ‘‘old’’ responses are reported as a function of the number of actually presented DRM list items that preceded the critical lure at test (again referred to as items preceding). It should be noted that correct memory for actually presented items decreased relative to Experiment 1 whereas false memory for all categories of unstudied items increased. This is a common consequence of the pressure to make

TABLE 2
Experiment 2

	Proportion of ‘‘Old’’ responses					
	Items preceding					
Item type	*0*	*1*	*2*	*3*	*4*	*5*
Critical lures	.84	.92	.85	.91	.84	.86
	(.04)	(.03)	(.04)	(.03)	(.04)	(.03)
Presented items	.81	.82	.81	.81	.82	.82
	(.02)	(.02)	(.01)	(.02)	(.01)	(.02)
Weakly related/unstudied items	.40	.38	.33	.48	.43	.39
	(.04)	(.04)	(.05)	(.05)	(.04)	(.03)
Unrelated/unstudied items	.24	.23	.21	.22	.24	.27
	(.05)	(.04)	(.03)	(.04)	(.05)	(.04)
Corrected critical lures	.60	.69	.64	.69	.60	.59
	(.06)	(.05)	(.06)	(.05)	(.07)	(.05)

Experiment 2: Mean probability of an ‘‘Old’’ response and Standard Error (SE) on the recognition test as a function of item type and of the number of studied items preceding the critical lure at test. Scores represent hits for presented items, and false alarms for critical lures, weakly related/unpresented items, and unrelated/unstudied items.

The final row reports corrected false recognition scores for critical lures, calculated as the false alarm rate to critical lures minus the false alarm rate to unrelated/unstudied items.

a speeded recognition decision and indicates that subjects were sensitive to the time pressure (Benjamin, 2001). Response times were also recorded—however there was no difference in response time as a function of experimental condition ($F < 1$).

To determine whether test-induced priming influences false recognition, mean proportions of ''old'' responses to critical lures were analysed using a one-way repeated measures ANOVA. There was no effect of items preceding, $F(5, 145) = 1.09$, $MSE = .06$, $p > .35$, indicating that false recognition was not influenced by the number of actually studied items preceding the critical lure at test. It is important to note that the anomalous finding from Experiment 1 (substantially lower false recognition when four actually studied items preceded the critical lure) did not replicate in the present experiment, suggesting that the prior result was an artefact of the low number of observations in each condition. To further emphasise this point, a 6 × 2 repeated measures ANOVA was conducted with items preceding as a within-subjects factor and experiment (1 or 2) as a between subjects factor. Importantly, there was no interaction between items preceding and experiment, $F(5, 315) = 1.02$, $MSE = .07$, $p > .45$.

As in Experiment 1, we investigated the effect of association strength for the 1 and 2 associates preceding conditions and found no effect of this variable. False recognition was unaffected by whether the critical lure was preceded by a weak associate(s) or a strong associate(s), $t(29) = 1.01$, $SE = .07$, $p > .30$, and $t(34) = 0.49$, $SE = .10$, $p > .60$, for the 1 and 2 associates preceding conditions, respectively.

To determine whether exposure to the critical lure influenced correct recognition of studied items, an analogous ANOVA was conducted on the mean proportions of ''old'' responses to actually studied items. There was no effect of items preceding, $F(5, 145) < 1$, indicating that correct recognition was not influenced by exposure to the critical lure at test. As in Experiment 1, it did not matter whether the critical lure appeared prior to the presentation of any studied item (0 condition) or following the presentation of all studied items (5 condition); correct memory in both conditions was virtually identical (.81 and .82 for the 0 and 5 conditions, respectively).

Analogous ANOVAs were also performed on the proportion of ''old'' responses to weakly related/unstudied items and to unrelated/unstudied items. There was a marginally significant effect of studied items preceding the critical lure on weakly related/unstudied items, $F(5, 145) = 2.22$, $MSE = .04$, $p < .06$, but this was not due to any systematic difference in the likelihood of false alarms to weakly related/unstudied items as a function of items preceding. Furthermore, there was no effect of items preceding on unrelated/unstudied items, $F(5, 145) < 1$.

The results of Experiment 2 are in accord with those of Experiment 1 and of Marsh et al. (2004): Test-induced priming does not influence false recognition of critical lures. Even under speeded conditions—where individuals have reduced access to elaborative retrieval strategies and might be expected to rely on familiarity (or gist, Brainerd & Reyna, 1998) and, therefore, to be more susceptible to false memory—the presentation of actually studied semantically related associates at test did not lead to an increase in false recognition for critical lures. The present results also suggest that recollection rejection did not counter the possible effects of test-induced priming in Experiment 1, given that the present results mirror those of Experiment 1 despite the fact that the time pressure prevented participants from using an elaborative retrieval strategy at test. Also confirming the finding in Experiment 1, we failed to observe any influence of the position of the critical lure at test on actually presented study items in Experiment 2.

Taken together, the results of the present experiments strongly suggest that the genesis of true and false memories in the DRM paradigm is likely at encoding, and confirm that the composition and sequencing of the recognition test has no influence on memory performance, even under conditions that should have promoted such an influence.

GENERAL DISCUSSION

The present experiments demonstrate that test-induced priming plays no role in the false recognition of critical lures in the DRM paradigm. In each of our experiments, false recognition of the critical lures was unaffected by the number of actually studied items preceding the critical lure at test.[4] Furthermore, our experiments demonstrate that correct recognition of actually studied items is not influenced by the position of exposure to critical lures at test. The implications of these findings both for the creation of false memories and for theories of false memory will be discussed in turn.

DRM illusions as an encoding phenomenon

Our results confirm the findings of Marsh et al. (2004) in demonstrating that test-induced priming plays no role in the creation of false memories. Marsh et al. demonstrated that test-induced priming did not occur when participants studied numerous DRM lists both auditorily and visually prior to a single batch recognition task. The present line of research extends this finding to strictly auditory presentation but with individual recognition tests immediately following each DRM list. Furthermore, test-induced priming was not observed, regardless of whether the recognition test occurred under speeded or unspeeded conditions. These results strongly suggest that false memories are formed at encoding, and that retrieval processes play no role in the endorsement of critical lures at recognition.

These results may seem somewhat surprising given that numerous manipulations have been reported which can lead to a reduction and/or enhancement in the occurrence of false memory, for example, serial learning (Read, 1996); short lists with few converging associates (Robinson & Roediger, 1997); repetition (Benjamin, 2001; Tussing & Greene, 1997). Critically, however, these other manipulations seem to affect false memory at the level of encoding (or they provide item-specific information that participants can later use to dissociate between true and false memories; Marsh et al., 2004). As noted by Marsh et al., the absence of test-induced priming on false recognition is entirely consistent with a growing literature suggesting that false memories are unaffected by test manipulations. For example, Gallo and colleagues (Gallo, Roberts, & Seamon, 1997; Gallo, Roediger, & McDermott, 2001) have shown that a post-study warning does not reduce false memory. Moreover, source monitoring, a process that has been shown to reduce memory errors across a variety of tasks—e.g., eyewitness suggestibility (Lindsay & Johnson, 1989); false fame effect (Multhaup, 1995)—does not reduce the prevalence of false memory (Multhaup & Conner, 2002); in some cases, it actually leads to an increase in false memory (Hicks & Marsh, 2001).

The notion that false memories are formed at encoding is also consistent with a number of other findings in the false memory literature. For example, Seamon, Lee, Toner, Wheeler, Goodkind, and Birch (2002) have demonstrated that participants often mistakenly rehearse the critical lure when required to rehearse overtly. Moreover, participants who spontaneously rehearsed the critical lure were far more likely to exhibit a later false memory than were participants who did not spontaneously rehearse the critical lure at study. In addition, although Gallo et al. (2001) have demonstrated that a post-study warning does not reduce false recognition, they have also demonstrated that a pre-study warning does substantially reduce (but does not eliminate) false recognition (see also Gallo et al., 1997; McDermott & Roediger, 1998). Taken together with the aforementioned research, the present results suggest that false memories are unaffected by test manipulations and are attributable to processes occurring at study. This is true even under conditions optimal for the production of false memories.

[4] We note that our conclusion—that test-induced priming does not influence false recognition—is a prediction consistent with the null hypothesis. Thus, we conducted a power analysis to determine the likelihood that we would observe a significant effect in the present study. In our Experiment 2, the largest difference in false recognition between conditions was .1. To have sufficient power to observe an effect size of .1 in the present study, we would have required 250 participants, which clearly would not be a feasible sample size for a study of this nature (to detect an effect size of .25 would have require 140 participants). We had sufficient power to detect an effect size of approximately .4 had such a difference existed. It is important to note that within-subject power analyses are often impractical and as such, may not be useful or practical in the present situation (see Howell, 1997). However, given that we used a speeded response manipulation and incorporated several levels of our independent variable, we are confident that our conclusion—that test-induced priming does not influence false recognition—is likely to be correct.

Activation/monitoring and fuzzy-trace theory

That re-exposure to items at test does not influence false recognition, particularly under speeded conditions, seems inconsistent with the monitoring component of activation/monitoring. The activation/monitoring account (Roediger et al., 2001) posits that false memories can occur either from conscious, elaborative processes, or from automatic spreading of activation to critical lures when DRM list words are processed. Critically, a monitoring process is also thought to influence the occurrence of false memory. This monitoring process is thought to be active during retrieval as individuals attempt to determine the source of item familiarity (e.g., participants should be able to determine the source of an item's activation level).

Consistent with this monitoring idea, Schacter and colleagues (e.g., Dodson & Schacter, 2001; Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999) have demonstrated that false memory is reduced by manipulations that increase the distinctiveness of study items. It has been suggested that when items are distinctive, this allows individuals to encode item-specific information that they can later use to discriminate between true and false memories (Marsh et al., 2004). Consequently, Marsh et al. suggested that the effects of activation manipulations at test (such as the present manipulation) could be obscured when participants are able to monitor for item-specific information.

If monitoring is disallowed or prevented, one would expect the prevalence of false memory to increase with increases in activation. In the present Experiment 2, however, the time pressure manipulation (which should eliminate access to the monitoring process) had no effect on false recognition of the critical lure as a function of how many studied items preceded the lure at test. Moreover, the finding that a post-study warning does little to reduce false memory also seems inconsistent with monitoring. If the purpose of the monitoring process is to distinguish between true and false memories as a function of activation level of the two memory types, then explicitly instructing participants to differentiate between true and false memories at test should reduce false memory. That it does not do so would again seem to suggest that monitoring plays little role in the prevalence of false memory.

Although the present results seem inconsistent with activation/monitoring, they are consistent with fuzzy trace theory (e.g., Brainerd & Reyna, 1996, 1998) for two important reasons. First, fuzzy trace theory stresses the role of relatively automatic retrieval mechanisms on memory tests relative to the monitoring component of activation/monitoring. Second, and more critically, fuzzy trace theory provides a reasonable explanation of why false recognition is unaffected by re-exposure to items at test. According to fuzzy trace theory, verbatim and gist traces have opposite effects on false memory. Specifically, verbatim retrieval should suppress false memory because it de-emphasises the meaning of items, whereas gist retrieval should enhance false memory because it emphasises the meaning of an item (see Brainerd & Reyna, 2002, for a review). When items are represented at test, they provide a good retrieval cue for both verbatim and gist traces (e.g., this strong retrieval cue for verbatim traces may be offset by the fact that each item also cues the meaning of the critical distractor). Consequently, the suppression of false memory associated with verbatim traces would counteract the enhancement of false memory that would be expected given gist retrieval. Although the present results do not mean that activation/monitoring is an incorrect account of false memory illusions, they appear to be more consistent with fuzzy trace theory.

Spreading activation

Our results (in combination with those of Marsh et al., 2004) have important ramifications for spreading activation theories. Specifically, they provide insight into the differential situations under which activation spreads. Spreading activation models state that when an item is processed, activation automatically spreads to semantic associates (see, e.g., Collins & Loftus, 1975). This is a common assumption in the false memory literature and is often used to account for DRM memory illusions (e.g., activation spreads to the critical lure each time a semantic associate is processed at encoding, leading to a high level of activation of the lure and subsequent confusion regarding the source of that activation at retrieval).

If activation spreads automatically during encoding, then it should be reasonable to assume that activation should automatically spread when items are processed at retrieval. However, the present results, as well as those of Marsh et al., suggest that this is not the case. Moreover, the present results also provide evidence that exposure to the critical lure plays little role in accurate memory performance at test. In accord with the notion that activation spreads to semantic associates when a word is processed, we hypothesised that exposure to the critical lure at test could also lead to a spreading of activation to the previously studied list items and that this could impact the recognition decision for subsequently encountered study items at test. We found no evidence of this, however.

Why would activation spread automatically during encoding but not during retrieval? One possibility is that each item is already at a peak level of activation during test. Although this argument could be made for the present results, where the lag between study and test was short, it is unlikely that items were at peak activation in the Marsh et al. study, where there was on average a long lag between an item's encoding and its subsequent test. Thus, we would argue that the absence of spreading activation at test is not attributable to items already being at a peak level of activation. Rather, it may be the case that the processes at encoding are conducive to spreading activation whereas the processes at retrieval are not.

If this is the case, then it is incorrect to assume that activation spreads to semantic associates any time an item is processed. Rather, it would seem as though the process that is being carried out on an item is a critical determinant as to whether activation will spread. For example, activation does appear to spread when individuals attempt to memorise items, or attempt to determine whether an item is a word or nonword (as has been demonstrated by numerous studies of semantic priming), but activation does not appear to spread when individuals make recognition decisions (as a caveat, see Underwood, 1965; also MacLeod & Nelson, 1976, where activation is assumed to spread during a continuous recognition task, whereby the recognition test also serves as the study session). At this point, however, this claim is very speculative in nature and further research will be required to determine whether there are differential situations under which activation spreads.

Conclusion

The present results, along with those of Marsh et al. (2004), suggest that test-induced priming does not influence false recognition of critical lures in the DRM paradigm, nor indeed does test-induced priming influence accurate recognition. These results are consistent with fuzzy trace theory but inconsistent with the monitoring process of activation/monitoring. Moreover, the present results suggest that spreading activation is not as automatic a process as has been believed, as spreading activation does not seem to occur during recognition tests. Although we cannot rule out the influence of response processes other than test-induced priming on false recognition, the present results suggest that the genesis of false memory is most likely to be entirely at encoding, contrary to the initial speculation of Roediger and McDermott (1995).

Manuscript received 30 April 2004
Manuscript accepted 27 September 2004
PrEview proof published online 21 June 2005

REFERENCES

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 941–947.

Benjamin, A. S., & Craik, F. I. M. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition*, *29*, 691–697.

Brainerd, C. J., & Reyna, V. F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology*, *32*, 467–478.

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*, 164–169.

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, *110*, 762–784.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing.*Psychological Review*, *82*, 407–428.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Dodd, M. D., & MacLeod, C. M. (2004). False recognition without intentional learning. *Psychonomic Bulletin & Review*, *11*, 137–142.

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

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.

Gallo, D. A., Roediger, H. L., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, *8*, 579–586.

Graves, R. E., & Bradley, R. (1991). Millisecond timing on the IBM PC/XT/AT and PS/2: A review of the options and corrections for the Graves and Bradley algorithm. *Behavior Research Methods, Instruments, & Computers*, *23*, 377–379.

Hancock, T. W., Hicks, J. L., Marsh, R. L., & Ritschel, L. (2003). Measuring the activation level of critical lures in the Deese-Roediger-McDermott paradigm. *American Journal of Psychology*, *116*, 1–14.

Hicks, J. L., & Marsh, R. L. (2001). False recognition occurs more frequently during source identification than during old–new recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 375–383.

Howell, D. C. (1997). *Statistical methods for psychology*. Belmont, CA: Duxbury Press.

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review*, *4*, 577–581.

Kellogg, R. T. (2001). Presentation modality and mode of recall in verbal false memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 913–919.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, *17*, 349–358.

MacLeod, C. M., & Nelson, T. O. (1976). A nonmonotonic lag function for false alarms to associates. *American Journal of Psychology*, *89*, 127–135.

Marsh, E. J., McDermott, K. B., & Roediger, H. L. (2004). Does test-induced priming play a role in the creation of false memories? *Memory*, *12*, 44–55.

McDermott, K. B., & Roediger, H. L. (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.

McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 507–520.

Multhaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology & Aging*, *10*, 492–497.

Multhaup, K. S., & Conner, C. A. (2002). The effects of considering nonlist sources on the Deese-Roediger-McDermott memory illusion. *Journal of Memory & Language*, *47*, 214–228.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385–408.

Read, J. D. (1996). From a passing thought to a false memory in 2 minutes: Confusing real and illusory events. *Psychonomic Bulletin & Review*, *3*, 105–111.

Robinson, K. J., & Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*, 231–237.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic 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.

Seamon, J. G., Lee, I. A., Toner, S. K., Wheeler, R. H., Goodkind, M. S., & Birch, A. D. (2002). Thinking of critical words during study is unnecessary for false memory in the DRM procedure. *Psychological Science*, *13*, 526–531.

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.

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 Psychology*, *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.