Available online at www.sciencedirect.com
ScienceDirect
Acta Psychologica 123 (2006) 240–260
acta psychologica
www.elsevier.com/locate/actpsy

An investigation of false memory in perceptual implicit tasks

Dawn M. McBride *, Jennifer H. Coane [1], Bascom A. Raulerson III [2]

Department of Psychology, Illinois State University, Campus Box 4620, Normal, IL 61790-4620, USA

Received 21 April 2005; received in revised form 12 January 2006; accepted 13 January 2006

Abstract

Reports of critical lure priming in perceptual implicit tasks [e.g., McKone, E., & Murphy, B. (2000). Implicit false memory: Effects of modality and multiple study presentations on long-lived semantic priming. *Journal of Memory and Language*, *43*, 89–109] using the Deese–Roediger–McDermott [Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814] procedure have suggested availability of the lexical form of lure items at study. Three experiments were conducted to further explore "false" implicit priming in perceptual tests. In Experiments 1 and 3, implicit and explicit stem completion tests were given in the DRM procedure with semantic lists; in Experiment 2, a graphemic response test was used in a similar design. For all experiments, explicit instructions resulted in reliable false memory, while implicit instructions resulted in priming for list items and no priming for lure items. Priming for lure items was evident for "test-aware" subjects only in Experiment 1 and in a combined analysis for all three experiments. These results establish boundary conditions for priming for critical lures and indicate that access to the lexical form of critical lures may not occur under incidental learning conditions when strong controls against explicit retrieval are implemented.

PsycINFO classification: 2343

Keywords: False memory; Implicit memory

---

* Corresponding author. Tel.: +1 309 438 7146; fax: +1 309 438 5789.
E-mail address: dmcbride@ilstu.edu (D.M. McBride).

[1] Present address: Department of Psychology, Washington University in Saint Louis.

[2] Present address: Department of Psychology, St. Louis University.

doi:10.1016/j.actpsy.2006.01.002

1. Introduction

Findings of experimentally induced false memories have been frequently reported in recent studies. Many of these studies have utilized a paradigm originally developed by Deese (1959) and recently popularized by Roediger and McDermott (1995). In these experiments subjects are shown or read a list of words (e.g., *dream*, *bed*, *blanket*, etc.) that are linked by a conceptual theme word (e.g., *sleep*) that is not presented. Subjects commonly exhibit rates of false memories for the theme word (also called a critical lure) in tests of recall or recognition that are similar to rates of accurate recall or recognition for the list items that were presented (Roediger & McDermott, 1995). This method is now known as the DRM procedure.

In attempts to understand how these false memories are created, researchers have investigated several variables that attenuate or enhance false memories as compared to true memories. For example, semantic study tasks with levels of processing manipulations increase both accurate memory for list items and false memory for lure items (Coane & McBride, 2002; Rhodes & Anastasi, 2000). This finding is not surprising, because the semantic study tasks are likely to strengthen the conceptual link between the list items and the critical item making them more likely to be remembered (or falsely remembered for the critical item) in the test.

A finding that is more surprising is the one reported by several researchers (Hancock, Hicks, Marsh, & Ritschel, 2003; McDermott, 1997; McKone & Murphy, 2000; Tajika, Neumann, Hamajima, & Iwahara, 2005; Tse & Neely, 2005) where implicit false memory was found in implicit tasks typically classified as perceptual in nature. In contrast with explicit memory tasks (such as recall and recognition), implicit tasks do not require intentional retrieval of studied items. Instead, subjects are asked to perform a task that involves studied items with no reference made to the study episode. Many implicit tasks have been classified as perceptual tasks, because the test cues involve a perceptual form of the studied items. For example, in a word stem completion task, three-letter word stems are provided as cues in the test for the subjects to complete with the first word they think of. In lexical decision, items are presented for speeded word/non-word judgments. Target (studied) item performance that exceeds baseline rates when the items have not been studied provides measurements of implicit memory in the tasks. Performance in perceptual implicit tasks (like stem completion and lexical decision where perceptually altered forms of the studied items are given as cues) tends to be affected by perceptual manipulations (e.g., change in study to test modality), but is less likely to be affected by conceptual manipulations (like level of processing). In addition, there should be little or no perceptual advantage for critical lure items (over other non-studied items) in perceptual implicit memory tests because they are linked to study items conceptually and are not seen in the study phase. Critical lures should only show priming effects in perceptual implicit tests if their lexical form is accessed when list items are studied.

Findings of false memory in perceptual implicit tasks have important theoretical implications. Evidence of priming for lures could be interpreted as evidence for long-term semantic priming. Semantic priming refers to the facilitation target items receive when the prime is semantically related (e.g., responses to *nurse* are faster and more accurate when the prime is *doctor* relative to when the prime is an unrelated word like *cat*). Whereas repetition priming (i.e., facilitation for items previously presented) tends to be long lasting, semantic priming tends to be very short-lived, and tends not to last across intervening items (Meyer, Schvaneveldt, & Ruddy, 1974). Semantic or associative priming is furthermore assumed to be automatic and does not require that the prime be consciously perceived (e.g., Lukatela & Turvey, 1994). In the standard presentation of DRM lists, when several items and extended periods of time occur between presentation of list items and the critical lure, semantic priming would not be expected to occur. Thus, results that indicate that the critical lure can be primed at long intervals would provide evidence that semantic priming may not be as short-lived as previously thought.

Furthermore, findings of false memory in implicit tasks can provide a test of one of the theories currently proposed as an explanation for false memory effects observed in the DRM paradigm. Roediger, Balota, and Watson (2001) have proposed that false memory effects created in the DRM procedure are due to a combination of two processes. They suggest that false memories for critical lures occur because (a) the lures are elicited (either consciously or automatically) due to semantic activation when the associated list items are studied and (b) the lures are mistakenly identified as studied when subjects make source errors in the memory test. This combination of processes is the activation-source account of false memory.

Tse and Neely (2005) propose that priming effects for critical lures in implicit memory tests provide a means of obtaining false memory only due to activation processes, because source processes should be absent in implicit tests due to the nature of implicit retrieval. They state that priming for lures in implicit tests should not be caused by source errors, because subjects are asked to complete the task without reference to the study episode. Therefore, the source of the responses is irrelevant. Using this logic, examinations of false implicit memory are examinations of false memory due to activation alone. The exception to this occurs when explicit retrieval is used in the implicit test (i.e., explicit contamination) and the source processes become important.

Tse and Neely (2005) evaluated false implicit memory using a lexical decision task, where subjects were asked to make word/non-word judgments for various types of test items. Some of the test items corresponded to DRM list items presented in a prior study session. Other test items were critical lures from the previously presented DRM lists. Test items were also included (list and lure items) for DRM lists that had not been studied. In four experiments, Tse and Neely found a reaction time advantage for responses to lures that were associated with studied lists as compared to lures associated to non-studied lists. They argued that their finding of critical lure priming even when no studied items were included in the lexical decision task prior to the critical lure item (experiment 4) provides strong evidence that priming for the lure can occur in perceptual tasks.

Although Tse and Neely (2005) found false implicit memory in lexical decisions, overall, studies of implicit false memory in lexical decision tasks have reported inconsistent results. Like Tse and Neely, Hancock et al. (2003) found significant priming for critical lures in lexical decision tasks in four experiments. However, the control items used by Hancock et al. may have artificially inflated estimates of critical lure priming. As argued by Tse and Neely, Hancock et al. may have used a critical lure comparison condition that did not produce appropriate baseline RTs (see Tse & Neely, 2005, for a discussion of this point). In addition, Meade, Watson, Balota, and Roediger (in preparation) found priming for critical lures in a lexical decision task only when the lures were presented early in the test session. When one or more intervening items were presented between the primes and the target, no priming for lures was found, indicating a rapid decay of priming effects.

Finally, Zeelenberg and Pecher (2002) found no priming for lures in a lexical decision task in their four experiments. McKone (2004) reported similar results for a lexical decision task. Therefore, the studies that used lexical decision as a measure of implicit false memory have provided mixed results.

Studies examining false implicit memory in word stem and fragment completion have provided more consistent results. In her study, McDermott (1997) examined false memory using the DRM paradigm with implicit word stem and fragment completion tests and an implicit word association test. Implicit word stem and fragment completion tests are typically classified as perceptual memory tests, because test cues include degraded forms of the words (three-letter stems or word fragments with letters replaced with dashes). In word association tasks, however, conceptually related items are presented as test cues. McDermott found significant false implicit memory for the critical lure items in the fragment completion and word association tests and a marginally significant false memory effect for the stem completion test, indicating that priming for the critical lures can occur in perceptual (as well as conceptual) implicit tasks.

One possible explanation of McDermott's (1997) findings of priming for the critical lures may be that subjects explicitly retrieved items in the implicit tests. As described above, explicit tests of memory (e.g., recognition, cued recall) typically result in false memory for the lures. If subjects noticed a connection between the test cues and the items they had previously studied, they may have attempted to recall those items and mistakenly retrieved the lures as they do in explicit tests of memory. In fact, subjects in McDermott's study were given 12–20 s to respond to the test cues, potentially allowing enough time for explicit retrieval to take place. In their study, McKone and Murphy (2000) attempted to test the explicit contamination explanation of McDermott's results for implicit word stem completion tasks. They conducted three experiments using the DRM and stronger controls for explicit contamination (speeded instructions and post-test questionnaires) to rule out explicit retrieval as an explanation of McDermott's findings of false implicit memory. Even with stronger controls for explicit contamination, their findings were similar to that of McDermott's: Significant priming for critical lures was found, providing additional evidence of false implicit memory in perceptual tests.

Despite better controls for explicit contamination, McKone and Murphy's (2000) results for implicit stem completion might still be caused by explicit retrieval by their subjects. For example, subjects were given a chance to rehearse the study items for 30 s before testing, thus making it more likely that subjects would notice a connection between test items and study items. In addition, subjects may have consciously rehearsed the critical lure, thus explicitly accessing the lexical form. Indeed, rehearsing the critical lure has been found to enhance later false recall (Seamon et al., 2002). Subjects were also given an implicit deadline to complete responses (i.e., they were told to spend 2 s on each test item), but no actual deadline was imposed and no timing prompt was given during testing to let subjects know when to move on to the next item.

Tajika et al. (2005) reported results similar to those of McKone and Murphy (2000). With incidental study tasks, subjects exhibited lure priming in an implicit stem completion test equal to priming for list items. However, like McKone and Murphy's results, Tajika et al.'s results may be due to explicit contamination in the implicit test and may not actually reflect priming. Their implicit test included stems from only 24 items: 12 list items, 4 critical lures, and 8 unstudied items. With so few unstudied items in the test, it is likely that subjects connected the study and test episodes and may have intentionally retrieved study items to complete the word stems. Previous studies have indicated that the proportion of related or studied items included in an implicit test can affect the probability that subjects engage in intentional retrieval strategies (e.g., Tse & Neely, 2005). Intentional retrieval of study items would also have brought to mind the critical lures, resulting in lure completion rates which are above the rates for unstudied items. Although Tajika et al. reported administering a post-test questionnaire to all their participants, as they did not report how it was worded it is difficult to assess whether such a questionnaire was effective in screening participants who engaged in intentional retrieval. In addition, Tajika et al. only included lures associated with studied lists in the test. Lure items may be unique in that they are items with a number of strongly associated items and baseline rates may be different for these items as compared to items in the study lists. Therefore, Tajika et al.'s unstudied completion rates (which did not include lures from unstudied lists) may not make an appropriate comparison for completion rates of lures associated with studied lists.

A comparison of false implicit memory results from studies with word stem and fragment completion tasks and studies with lexical decision shows that more consistent priming effects for lures are reported for the completion tasks. However, the consistent results could be due to consistent explicit contamination in these completion tasks (lexical decision tasks are less likely than stem and fragment completion tests to be contaminated by explicit retrieval, because the task is timed and tends to be much easier for subjects to complete). Therefore, the current study was designed to further investigate false implicit memory effects found in perceptual completion tasks described above with stronger controls for explicit contamination compared to prior studies. Perceptual memory tests were used in an attempt to support and generalize findings of false implicit memory in perceptual tests reported by McDermott (1997), McKone and Murphy (2000), and Tajika et al. (2005), as these results might have been contaminated by explicit retrieval attempts.

Two tests typically classified as perceptual tasks (word stem completion and graphemic cued response) were used with implicit and explicit instructions in the DRM paradigm. To examine and reduce possible explicit contamination in the implicit tests, several checks were included in the method of the experiments. First, a level of processing manipulation was employed at study in the first two experiments. Explicit memory tasks typically show an advantage for items studied in a semantic task over a graphemic task, while implicit tasks rarely result in significant levels of processing effects (Roediger & McDermott, 1993). In all experiments, participants encoded items under incidental learning instructions to reduce the probability of explicitly accessing the lexical form of the critical lure during rehearsal and to minimize the chances that subjects in the implicit test conditions would engage in intentional retrieval. In addition, a time limit per trial (2–3 s, depending on the task) was imposed in the implicit tests to discourage use of explicit strategies. Test cues also included a large number of unstudied items other than the critical lures (68% fillers in Experiments 1 and 2; 75% fillers in Experiment 3) to hide the nature of the implicit tests and to avoid potential strategic changes on the part of participants due to high relatedness proportion. Lastly, subjects received a post-test questionnaire to determine if they had used explicit retrieval and/or become aware of the connection to the study items in the implicit tests. Subjects who indicated that they had used an explicit retrieval strategy were eliminated from the analyses. Subjects were classified as "test-aware" if they reported in the post-test questionnaire that they recognized a connection between study and test items and as "test-unaware" if they did not. Richardson-Klavehn and Gardiner (1996) have suggested that subjects who are aware of the study-test connection in implicit memory tests may use different retrieval processes than those who remain unaware, even in the absence of intentional retrieval. The analyses were conducted with awareness included as a between-subjects factor to test this possibility in the current study.

2. Experiment 1

Experiment 1 was designed in an attempt to replicate and extend previous findings of implicit false memory for word stem completion. The DRM paradigm was used with implicit word stem completion and explicit cued recall tasks. Level of processing was manipulated at study to test for explicit contamination in the implicit tests.

2.1. Method

2.1.1. Participants

One hundred and seven participants were recruited from Illinois State University's psychology department subject pool. All were over the age of 18 and native speakers of American English. Fifty-eight participants were assigned to the implicit test condition and 49 were assigned to the explicit condition. In this and in the other two experiments, more subjects were assigned to the implicit test condition, as this was the primary condition of interest and some subjects' data were deleted in this test condition based on responses to the post-test questionnaire.

2.1.2. Design and materials

A 3 (study condition: semantic vs. graphemic level of processing vs. unstudied) × 2 (item type: list items vs. critical lure) × 2 (retrieval instructions: implicit vs. explicit) mixed factorial design was used. Study condition and item type were manipulated within subjects; retrieval instructions were manipulated between subjects.

Twenty-two DRM lists of semantic associates from the Stadler, Roediger, and McDermott (1999) study were initially tested in a norming session to select critical lures with a sufficiently low baseline completion rate to allow detection of priming effects.[3] The lists to be normed were selected according to several criteria. All words (list items and critical lures) had to be at least four letters in length. Words all had a unique three-letter stem. After these criteria were applied, remaining lists were normed in a separate session to determine baseline rates for list and lure items. With these constraints in place, the nine items comprising each list were not always the nine highest associates of the critical lure (based on the Stadler et al. norms), but all were among the 15 highest associates. Baseline data from 45 subjects were collected and any lists with critical lure completion rates of .50 or greater were eliminated. Fifteen lists remained with an average critical lure completion

---

[3] Prior studies conducted in our laboratory using stem completion (Coane & McBride, 2002) and fragment completion (Coane & McBride, unpublished data) indicated that when the baseline completion rates were very high (i.e., .40 or greater), detection of priming effects was quite unlikely. To minimize the risk of ceiling effects for completion rates, we determined that a baseline completion rate for critical lures of approximately .20 might allow detection of priming effects, while still providing evidence that the lures were spontaneously generated in a baseline task (and were thus not extremely low frequency items).

rate of .20 (SD = .17).[4] Although the majority of DRM studies have used longer lists (i.e., 12–16 items), pilot studies in our laboratory indicated that shorter lists of nine items reliably elicited false memories to non-presented critical lures under explicit testing conditions. Robinson and Roediger (1997) also found significant rates of false recall (approximately .20 in two experiments) following the study of nine item lists. Furthermore, the stem completion task imposed constraints that limited the number of available items. As will be shown below, results on the explicit tests provide converging evidence of false memories even with short lists.

The 15 lists were divided into three sets of five lists each such that the average critical lure baseline completion rate from the norming session was similar for all three sets. The average baseline completion rates for the critical lures for the three sets were .206, .208, and .212. Each set was assigned to the semantic processing condition, the graphemic processing condition, or to the unstudied condition with an approximately equal number of times across participants. Thus, each participant provided five data points for each condition (studied—semantic processing, studied—graphemic processing, non-studied baseline). In the semantic processing task, subjects were asked to rate the pleasantness of each word on a scale of 1–7; in the graphemic processing task, they were to indicate the number of vowels contained in each word. The level of processing tasks was indicated by a "!" or "*" shown immediately before each item. A "!" indicated the subject was to perform the semantic processing task, and a "*" indicated they were to complete the graphemic task. All items within a list were processed under the same encoding instructions.

The study session was administered by computer. Four practice items preceded the presentation of 10 lists of nine items each. No cues indicated the breaks between lists. Participant's responses to the levels-of-processing task were recorded in booklets during the group testing sessions. At the beginning of the session, each participant was provided a booklet to record their answers. Each booklet contained 90 numbered lines (and four for the practice items) containing a "!" and a "*" (which participants were instructed to circle, depending upon which cue preceded the word) and digits from 1 to 7. Subjects circled the number corresponding to their response. Cue symbol definitions were printed at the top of each page in the booklets.

The test items included the 15 critical lures from the lists (10 studied, 5 unstudied) as well as list items selected from the 1st, 5th, and 9th serial position on each list. Three-letter stems were typed in random order in two columns, in booklets provided to the participants after the study session. Subjects were instructed to complete each stem by writing their responses in a blank space next to each stem. The order in which the stems were presented was the same for all participants. All stems had at least one alternative completion other than the studied item. Filler cues (stems that could not be completed with any target list or critical lure words from the list sets) were intermixed with targets to reduce explicit retrieval attempts in the implicit task. The test thus included 124 items (15 lures, 45 list items, and 64 fillers) of which only 30 had actually been studied. Items from all of the lists in each of the three list sets (described above) were included in the test (i.e., five lures and 15 list items from the semantic, graphemic, and unstudied lists). The booklet given to participants in the

---

[4] It is important to note that, although the baseline completion rates were relatively low, the lists selected following the baseline study were not weaker lists in regards to probability of eliciting the critical lure in explicit tests. In fact, according to the Stadler et al. (1999) norms, several of the lists utilized in the experiments were among the stronger lists (in fact, the majority of the lists used were ranked among the 16 lists with highest rates of false recall in the Stadler et al. norms).

explicit condition was identical to that used in the implicit test, except for the instructions provided on the cover sheet. At the end of each booklet given to subjects in the implicit test condition, the final questionnaire was included. A written debriefing statement was also included in all booklets.

2.1.3. Procedure

The majority of participants were run in groups, ranging in size from 3 to 18 people in each group. The average group size consisted of nine participants. Eleven participants were tested individually. Subjects were told that the study was designed to test how people process and respond to words. No mention was made of the tests in either the explicit or the implicit test conditions. Each participant was initially given a booklet for the study session, and instructions were read aloud. The stimuli were projected onto a screen in the testing room. The orienting cue remained on the screen for 2000 ms. Each word remained on the screen for 3000 ms. A blank screen, lasting 1500 ms, preceded the appearance of the next cue. Subjects recorded study responses in the study booklets. After all the lists had been presented, the experimenter handed out test booklets, according to the condition being tested (implicit or explicit). Participants were instructed not to open the booklets until they were told to do so. The experimenter read the instructions on the screen (which were also printed on the cover of each booklet with an example) out loud.

In the implicit test condition, participants were instructed to respond by writing down the first word they could think of to complete a stem. No reference was made to the study session. Instead, subjects were reminded that the task was designed to examine how people process and respond to words. They were further told they had to move to the next stem within 2 s (at which point they would hear a tone from the computer). The response interval was imposed to minimize the risk of explicit retrieval. Subjects were informed that they could move on to the next cue before the prompt if they were able to respond before the deadline, but that they should not stay on a cue after they heard the tone, regardless of whether they had been able to complete the stem. The importance of using the first word they could think of was emphasized, and subjects were told it was acceptable to leave an item blank if they were unable to come up with a response within the time allowed. Participants were urged to respond as quickly as possible and not to go back to any stems they might have previously left blank.

As a further control against explicit retrieval attempts, a final questionnaire was administered to all subjects in the implicit condition asking whether they had noticed that some of the stems could be completed with items they had studied, and, if so, whether they had deliberately attempted to recall items from the study session. Data from subjects who answered *yes* to the second question were excluded from further analysis. Subjects who responded *yes* only to the first question were classified as “test-aware,” whereas those who responded *no* to both questions were classified as “test-unaware.” In analyses of data from the implicit tests, test awareness (“test-aware” vs. “test-unaware”) was included as a between-subjects factor, to assess whether being aware of a relationship between the encoding session and the test interacted with any of the other variables.

In the explicit test, participants were told to try to recall as many words as possible from the study session. No response deadline was imposed in the explicit test, and subjects were strongly discouraged from guessing.

2.2. Results and discussion

2.2.1. Implicit task

Based on responses provided to post test questionnaires after the implicit test, two participants' data were excluded, and two other subjects' data were eliminated for not responding to the questionnaire, leaving 54 data sets for analysis. The stem completion data were submitted to a 2 (item type) × 2 (levels of processing) repeated measures ANOVA. Table 1 provides average completion rates for list items and critical lures. Significantly more list items ($M = .29$) were used to complete stems than critical lures ($M = .18$), $F(1, 52) = 25.24, p < .001$, MSe = .02.

The level of processing manipulation did not produce a significant effect on response rates, $F(1, 52) < 1.0$. There was also no significant interaction of item type and level of processing, $F(1, 52) < 1.0$.

Using target responses to lists from the unstudied condition as a baseline measure, paired samples t-tests compared response rates to studied lists and related lures with unstudied lists and related lures to determine levels of priming. Priming was significant for list items studied under both levels of processing conditions, $t(53) = 5.98$, $p < .001$ for semantically encoded lists and $t(53) = 6.47$, $p < .001$ for graphemically encoded lists. Neither critical lures related to semantically studied lists nor lures related to graphemically studied lists produced significant levels of semantic priming, $t(53) = 1.10$, $p = .28$ and $t(53) = .69$, $p = .49$, respectively. Therefore, no significant semantic priming was found for critical lures, although a numerical trend appeared to be present for critical lures related to lists encoded with the semantic study task.

Finally, we included test awareness as a between-subjects factor in the analysis. Thirty-four subjects were classified as "test-aware" and 20 as "test-unaware." The main effect of awareness was significant, $F(1, 52) = 6.94$, $p = .011$, MSe = .13, indicating that subjects who were "test-aware" showed greater priming overall ($M = .26$) than those who were "test-unaware" ($M = .21$). The awareness by item type interaction was marginally significant, $F(1, 52) = 2.97$, $p = .091$, MSe = .05. Planned comparisons indicated that aware subjects showed significant lure priming for semantically processed lists, $t(33) = 2.69$, $p = .011$, and marginally significant lure priming for graphemically encoded lists, $t(33) = 1.85$, $p = .07$. Subjects classified as unaware showed no priming for lures, ps > .13. Interestingly, the difference appears to be driven by different baseline completion rates to unstudied lures: Subjects classified as aware had lower baselines ($M = .13$) than subjects classified as unaware ($M = .22$). In this case awareness of the study-test connection may have caused

Table 1
Mean stem completion rates for list items and critical lures as a function of processing at encoding (standard deviations in parentheses) for Experiment 1

Item type	Level of processing		
	Semantic	Graphemic	Unstudied
Implicit test			
List items	.28 (.11)	.29 (.10)	.17 (.10)
Critical lure	.21 (.19)	.15 (.14)	.17 (.17)
Explicit test			
List items	.37 (.19)	.11 (.12)	.01 (.03)
Critical lure	.23 (.18)	.05 (.10)	.03 (.07)

subjects to adopt a more conservative response criterion. This could indicate that subjects who were aware that studied items were being tested were responding to the stems based on awareness that an item had been studied (even if they were not intentionally retrieving the items), and this more conservative response strategy resulted in fewer target responses. The effect of awareness on priming will be discussed further in Section 5.

2.2.2. Explicit task

Two subjects' data in the explicit test condition were omitted from analyses due to experimenter error and for failure to follow instructions, leaving 47 data sets for the analysis. Another 2 × 2 repeated measures ANOVA was conducted on responses provided in the explicit task instructions condition. Mean rates of recall for list items and lures are presented in Table 1. The main effect of item type was significant, with list items ($M = .24$) being recalled significantly more often than lures ($M = .14$), $F(1,46) = 27.52$, $p < .001$, MSe = .02. The levels of processing the main effect were also significant, with semantic processing leading to higher rates of recall ($M = .30$) relative to graphemic processing ($M = .08$), $F(1,46) = 124.79$, $p < .001$, MSe = .02. Finally, a marginally significant interaction effect of levels of processing and item type was found, $F(1,46) = 3.94$, $p = .053$, MSe = .02. Paired samples t-tests revealed that responses in the semantic encoding condition were significantly greater than in the graphemic encoding condition for both list and critical lure items, $t(46) = 8.97$, $p < .001$ and $t(46) = 6.38$, $p < .001$, respectively, although the levels of processing effect was larger or list items than for lures. This result for lures is likely due to the lack of false memory found for graphemically encoded lists in this experiment (see below).

List and lure recall rates for unstudied lists were very low, as is expected in explicit tests where subjects are not relying on guessing to generate responses (see Table 1 for unstudied list recall rates). Using the recall rates for lures related to unstudied lists as a baseline measure, paired samples t-tests compared response rates to studied list items and related lures with unstudied list items and related lures to verify false memory effects in the explicit test. False explicit memory was significant only for semantically encoded lists, $t(46) = 7.42$, $p < .001$. Semantic encoding of lists resulted in a false memory rate of about .20 (see Table 1), which is comparable to rates reported by Robinson and Roediger (1997) for recall of nine-item DRM lists and McDermott (1997) for recall of 12-item DRM lists (her explicit comparison condition for three implicit tests). False memory was not found for lists studied with the graphemic task, $t(46) = 1.35$, $p = .183$.

In conclusion, results from Experiment 1 indicate that priming was evident for lures only when subjects were aware of the connection between study and test, whereas significant priming for list items was found in aware and unaware subjects. In addition, based on the comparison of the implicit and explicit test results, it appears that explicit contamination was minimal in the implicit test condition.

3. Experiment 2

Experiment 2 was conducted to extend the results of Experiment 1 to another implicit perceptual task that is similar to stem completion in terms of the type of response and performance measure. The DRM procedure was used with implicit and explicit graphemic cued response tests. This task was originally developed by Blaxton (1989) and was modified to fit the DRM procedure. Participants were presented with orthographically similar words to the target items and asked to generate an item (a studied item for the explicit test or the first one they think of for the implicit test) that is similar in sound or appearance to the cue item.

3.1. Method

3.1.1. Participants

Participants for Experiment 2 included 111 undergraduates recruited from the same population as in Experiment 1, who had not participated in the previous experiment. Forty-nine subjects were randomly assigned to the explicit test condition, and 62 were assigned to the implicit test condition.

3.1.2. Design and materials

The 15 lists from Experiment 1 were used as stimuli in Experiment 2. Data from the same norming session described earlier indicated that the average baseline completion rate for critical lures was .09 (SD = .08). The same sets developed for Experiment 1 were used to counterbalance study task (semantic, graphemic, unstudied) assignments. Approximately the same number of subjects received each assignment of list set to study task condition.

Study and test stimuli were presented in the same manner as in Experiment 1, except that graphemic word items were presented as test cues in the implicit and explicit tests. The cues were developed by changing at least one letter in each of the target items to produce English words that were similar in sound, appearance, or both to the targets (e.g., *apple* or *maple* would be acceptable responses to the cue *ample*). The cues and their specific targets were not semantically related, although some of the cues could be interpreted as being semantically associated to other targets. All cues could yield at least one alternative response other than the intended target.

3.1.3. Procedure

The procedure for Experiment 2 was identical to that used in Experiment 1, except for the following: (a) subjects received the graphemic cued response test cues in the test phase and (b) the time limit to respond was 3 s in the implicit test. In the tests, subjects were instructed to write (next to the word provided) another word that looked or sounded like the cue item that either was studied earlier (explicit test condition) or was the first word they thought of that fit the cue (implicit test condition).

3.2. Results and discussion

3.2.1. Implicit task

Data from 12 subjects were omitted from analyses because of failure to follow instructions (by generating semantically related responses or non-words as responses to the cues) or based on their responses to the post-test questionnaire, leaving 50 data sets for analysis. Data were analyzed in a 2 (level of processing) × 2 (item type) repeated measures ANOVA. Table 2 contains means and standard deviations for target completion rates. As in the previous experiment, the mean completion rate for list item targets ($M = .18$) was significantly higher than the mean completion rate for lure items ($M = .05$), $F(1,48) = 73.01$, $p < .001$, MSe = .009.

Table 2
Mean graphemic cued response target completions as a function of item type and levels-of-processing at encoding (standard deviations in parentheses) for Experiment 2

Item type	Level of processing		
	Semantic	Graphemic	Unstudied
Implicit test			
List items	.17 (.11)	.18 (.12)	.12 (.08)
Critical lure	.05 (.11)	.05 (.10)	.04 (.09)
Explicit test			
List items	.27 (.19)	.09 (.09)	.01 (.03)
Critical lure	.11 (.16)	.03 (.08)	.01 (.05)

No main effect of level of processing was observed, $F(1,48) < 1.0$, which indicates that explicit contamination was not a significant factor in these data. The interaction between level of processing and item type was also not significant, $F(1,48) < 1$.

To assess the effects of repetition and semantic priming, responses to studied list items and related critical lures were compared to the non-studied baseline items. Significant priming was observed for both semantically and graphemically encoded items, $t(49) = 2.56$, $p = .014$ and $t(49) = 1.04$, $p = .005$, respectively. However, consistent with results from Experiment 1, no priming for critical lures was observed, regardless of encoding manipulations (both $ps > .30$).

Finally, the main effect of awareness was not significant ($p = .41$), nor did this factor interact with any other variables (all $ps > .05$). Of the 50 subjects, 31 were "test-aware" and 19 were "test-unaware."

3.2.2. Explicit task

Data from five subjects were excluded from the analysis because of failure to follow instructions. The data from the remaining 44 subjects were submitted to a 2 × 2 ANOVA with level of processing and item type as factors. A main effect of item type was observed, with list items ($M = .18$) recalled at a higher rate than critical lures ($M = .07$), $F(1,43) = 32.00$, $p < .001$, MSe = .01. The effect of level of processing was also significant, as semantic encoding ($M = .17$) yielded higher cued recall rates than graphemic encoding ($M = .09$), $F(1,43) = 55.47$, $p < .001$, MSe = .01. Finally, the interaction between processing instruction and item type was also significant, $F(1,43) = 5.14$, $p = .028$, MSe = .02.[5]

The recall rates for unstudied list and lure items were low for the explicit test (see Table 2). These rates are consistent with subjects' adherence to instructions not to guess in the explicit test. To verify false memory effects in this test, recall rates for lures related to

---

[5] A separate analysis was performed on the graphemic cued response data adopting a more liberal response criterion. As the cues on the test were developed by changing target items, and as several items in DRM lists share orthography and phonology to a certain extent, certain cues could elicit more than one response (e.g., *sing* from the *music* list and the critical lure *king* could have both been valid answers to the given cue *ring*). Furthermore, given the task instructions, it was possible that participants gave as responses studied list items that were not the intended targets, but had been studied. Thus, we ran the analyses including all list items given by subjects instead of including only the target items. Importantly, these additional analyses were performed on raw numbers, as it was not possible to generate proportions due to the variability in responses across subjects. The pattern of results was identical in the more liberal data as in the original data for both implicit and explicit tasks.

unstudied lists were compared with rates for lures related to studied lists. Significant false memory was found only for lures related to semantically encoded lists, $t(43) = 3.91$, $p < .001$. No significant false memory was found for lists studied with the graphemic task, $t(43) = 1.16$, $p = .253$.

Results from Experiment 2 are thus consistent with those from Experiment 1; no priming for lures was found although significant rates of false explicit memory were present, at least when the lists were encoded under semantic processing conditions.

4. Experiment 3

Experiment 3 was conducted to (1) attempt to replicate the results for stem completion found in Experiment 1 with a design that should increase the size of any false memory effects that might exist for this task and (2) examine implicit false memory under conditions that more closely mirror those used by McKone and Murphy (2000). To achieve a more similar design to their study and to increase the likelihood of finding an implicit false memory effect, more items per DRM list were included at study (study lists included 15 semantically associated items as compared to nine items per list in Experiments 1 and 2) and more lists were included for each condition (eight lists were used as compared to five lists per condition in Experiments 1 and 2). Experiment 3 focused on the semantic study condition only, as this was the condition that showed the most promise in producing a false memory effect in the implicit test based on the results reported in Experiment 1 for stem completion and pilot studies conducted prior to Experiments 1 and 2.

4.1. Method

4.1.1. Participants

Ninety-three undergraduate students from the same population of the prior experiments participated in Experiment 3. None had participated in either of the previous studies. Fifty-four participants were assigned to the implicit test condition and 39 were assigned to the explicit test condition.

4.1.2. Design and materials

The same materials used in Experiments 1 and 2 were used in Experiment 3 with the following exceptions: One list (the *cold* list) was removed, and two lists were added (*sleep* and *flag*), bringing the total number of lists to 16. All lists included 15 items. In most cases, these were the 15 associates listed in Stadler et al. (1999). In a few cases in which a list item had the same three-letter stem as a critical lure, this item was replaced with an associate from the Russell and Jenkins (1954) norms or the Nelson, McEvoy, and Schreiber (1998) norms.

The study lists for Experiment 3 included several three-letter words. The lists also included several words that shared the initial three-letter word stem with other items on the same or another list. None of these items were used as targets in the stem completion task; however, in order to use longer lists, inclusion of these items was necessary. The target list items for the test were selected according to the following constraints: (1) the list item was at least four letters in length; (2) no other list (or critical lure) item shared the same stem; (3) the stem was not an existing word (i.e., *button* was not selected as a target because *but* is a word); and (4) at least one other word (besides the target) could be used to complete the stem (in most cases, at least five common words could be used for possible completions).[6] Given these constraints, the target list items did not always come from the same serial position across all lists. However, as much as possible, one target was selected from the first five items in a list, one from the middle five, and one from the last five. The 16 lists were divided into two sets of eight lists. Each participant studied one set while the other set served as a non-studied baseline. Each list was in the studied and unstudied conditions for about half of the subjects.

The test booklets included the stems from three list items and critical lures from the eight studied lists, three list items and critical lures from the eight non-studied lists, and 64 filler stems that could not be completed with any items from any list. Thus, the proportion of primed (or studied) items, including the critical lures, was 25%.

4.1.3. Procedure

The procedure was the same as in the previous experiments, except that all items were encoded under semantic processing instructions and each list included 15 items instead of nine. In comparison to Experiments 1 and 2, the study session was slightly longer, as it included 120 items instead of 90.

4.2. Results and discussion

4.2.1. Implicit task

Data from seven subjects were omitted from analyses based on their responses to the post-test questionnaire (i.e., they reported using explicit retrieval in the implicit test), leaving 47 data sets for analysis. Data were analyzed in a 2 (study condition) × 2 (item type) repeated measures ANOVA. Table 3 contains means and standard deviations for target completion rates. In contrast to the previous experiments, the mean completion rate for list item targets ($M = .16$) was significantly lower than the mean completion rate for lure items ($M = .21$), $F(1,45) = 11.34$, $p = .002$, MSe = .133. The main effect of the study condition was also significant (more completions for studied items than unstudied items), $F(1,45) = 26.43$, $p < .001$, MSe = .242, as was the interaction of item type and study condition, $F(1,45) = 10.56$, $p = .002$, MSe = .098.

To assess the effects of repetition and semantic priming, responses to studied list items and related critical lures were compared to the non-studied baseline items. Significant priming was again observed for studied list items, $t(46) = 7.90$, $p < .001$. However, consistent with results from Experiments 1 and 2, no priming for critical lures was observed, $t(46) = 1.14$, $p = .262$. The unstudied baseline rate for lures ($M = .20$) was higher than that for list items ($M = .10$) in Experiment 3. However, as compared to the baseline rates for Experiment 1 ($M = .17$ for both item types), the rates for Experiment 3 represent an unusually low baseline rate for list items (rather than a much higher rate for lure items). These findings indicate consistency across experiments for the lack of priming effects for lures. Additional evidence of the overall consistency of the results found in all three experiments is addressed in Section 5.

---

[6] In one case, for the *mountain* list, the target *range* was selected, although the stem (*ran-*) did form an existing word. This was required because other possible study items in the list had already been excluded based on the other constraints listed.

Table 3
Mean stem completion rates for list items and critical lures as a function of study condition (standard deviations in parentheses) for Experiment 3

Item type	Study condition	
	Studied (semantic)	Unstudied
Implicit test		
List items	.22 (.09)	.10 (.06)
Critical lure	.22 (.15)	.20 (.10)
Explicit test		
List items	.22 (.09)	.02 (.03)
Critical lure	.27 (.15)	.02 (.04)

Finally, the main effect of awareness was not significant ($p = .31$), nor did this factor interact with any other variable (all ps > .18). Of the 47 subjects, 27 were classified as "test-aware" and 20 were classified as "test-unaware."

4.2.2. Explicit task

Data from the 39 subjects assigned to the explicit test condition were submitted to a 2 × 2 ANOVA with study condition and item type as factors. A marginally significant main effect of item type was observed, with lures ($M = .14$) recalled at a slightly higher rate than list items ($M = .12$), $F(1,37) = 2.98$, $p = .093$, MSe = .017. The effect of study condition was also significant; studied lists ($M = .25$) yielded higher cued recall rates than unstudied lists ($M = .02$), $F(1,37) = 219.39$, $p < .001$, MSe = 1.99. Finally, there was a marginally significant interaction effect for study condition and item type, $F(1,37) = 3.27$, $p = .079$, MSe = .019. As can be seen in Table 3, unstudied rates were identical for list and lure items, but lures from studied lists ($M = .27$) were recalled more often than studied list items ($M = .22$).

As in the previous experiments, recall rates for unstudied list and lure items were low for the explicit test (see Table 3). These rates are consistent with subjects' adherence to instructions not to guess in the explicit test. To verify false memory effects in this test, recall rates for lures related to unstudied lists were compared with rates for lures related to studied lists. Significant false memory was found for lures related to studied lists, $t(38) = 11.18$, $p < .001$.

Results from Experiments 3 are thus consistent with those from the previous two experiments. No false implicit memory was found for lures associated to studied lists even though significant rates of false explicit memory were present.

5. General discussion

Inconsistent with the findings reported by McDermott (1997), McKone and Murphy (2000), and Tajika et al. (2005) for stem and fragment completion tasks, but consistent with some findings reported for lexical decision tasks (Zeelenberg & Pecher, 2002), false implicit memory was not significant in the perceptual tasks used in the current study. Except for the marginal effect found for aware subjects in Experiment 1, significant priming for lures was not observed in either the implicit stem completion or the implicit graphemic response tasks. As expected, significant false memory for critical items was found in the explicit versions of these tasks, along with implicit and explicit memory for the list items.

One explanation for the lack of priming found for lure items could be the size of these effects. Priming of the lures in perceptual implicit memory tests could require more activation of the lure items than explicit false memory for the lures, because semantic associations appear to increase memory in explicit tests, but do not appear to increase memory much in implicit tests (Roediger & McDermott, 1993). Therefore, priming of critical lures in perceptual implicit tests may only be due to lexical activation of the critical lure at study, while recall of the lures in explicit tests may be due to both lexical activation of the lure and semantic associations between the lure and the list items being re-activated at test. Effect sizes for lure priming were estimated from statistics reported by McDermott (1997, Experiment 3) and by McKone and Murphy (2000, Experiments 1–3) for stem completion. Cohen's d values ranged from .058 to .260 in these studies. These values fall in the range defined by Cohen (1988) as small effects. Using the G*Power program (Erdfelder, Faul, & Buchner, 1996) and a small effect size ($d = .20$), power of the statistical tests for the current study was estimated at approximately .18.

To further increase the power to detect false implicit memory effects, an ANOVA was conducted on the combined implicit test data from all three experiments (semantic study and unstudied conditions only) with a total of 151 subjects. In this test, all effects were found to be significant, including the study condition by item type interaction, $F(1,150) = 11.44$, $p = .001$, MSe = .144. Paired samples t-tests indicated a significant priming effect for list items, $t(150) = 8.90$, $p < .001$, while the priming effect for lures was marginally significant, $t(150) = 1.80$, $p = .073$. However, when the same tests were conducted separately for "test-aware" and "test-unaware" subjects, the priming effect for lures was significant only for the "test-aware" subjects, $t(91) = 2.94$, $p = .004$, and not for the "test-unaware" subjects, $t(58) = -.56$, $p = .580$. These results are consistent with those found for "test-aware" subjects in Experiment 1 and indicate a possible constraint on false implicit memory effects such that they may occur only when subjects are aware of the connection between study and test items.

According to Richardson-Klavehn and Gardiner (1996), being "test-aware" may indicate use of an involuntary form of conscious memory. If this is the case for false implicit memory effects found in previous studies (e.g., McKone & Murphy, 2000), the effect may not be based on automatic forms of memory, as previously proposed. It is also possible that subjects who were "test-aware" used explicit retrieval on some trials, but did not do so often enough to warrant a *yes* response to using explicit retrieval on the post-test questionnaire. If this occurred, the effect across experiments reported above for "test-aware" subjects may also include small amounts of explicit contamination that were weak enough to only emerge in the overall analysis. It does appear, however, that if false implicit memory effects exist for "test-aware" subjects, the effect size must be particularly small, because test awareness only interacted with other factors in Experiment 1 and in the overall analysis.

Low effect sizes may be caused by methodological details in the DRM procedure. Due to the constraints of the cues needed for the stem completion and graphemic response tests, only a small number of lists can be used. In the current study, five lists were assigned to each level of processing condition in Experiments 1 and 2, and eight lists were assigned to each study condition (studied and unstudied) in Experiment 3. In their experiments, McKone and Murphy (2000) included eight lists per test condition (comparable to our Experiment 3), while McDermott (1997) used 12 in her experiments. Tajika et al. (2005) used only four lists per condition. In Experiments 1 and 2 of the current study only nine study items were included per list (a number sufficient to produce false memory in the explicit tests, see also Robinson & Roediger, 1997), while McKone and Murphy's and Tajika et al.'s lists contained 15 items and McDermott's contained 12 items. However, in Experiment 3 of the current study, we increased list length to 15 items per list, but still did not detect a false implicit memory effect. Incidental study conditions were used in the current experiments, while McDermott's and McKone and Murphy's studies informed subjects of a later memory test (making it more likely that their subjects were "test-aware"). These methodological differences may have resulted in small differences in activation for the lure items causing significant priming in their studies and no priming in the current study. Tajika et al. also used incidental study conditions, but the small number of unstudied items in their implicit test could have contributed to explicit retrieval or test awareness by the subjects.

There is always the possibility that explicit contamination in the implicit test will increase the chance of finding lure priming. The current study employed the most stringent controls for reducing and detecting explicit contamination. These controls included the following: (a) A level of processing study manipulation was included in each experiment. Results indicated no level of processing effect in any of the implicit tests, a result that is consistent with a lack of explicit contamination. A level of processing effect was found in the explicit tests for all experiments, as is predicted in explicit tests. (b) Subjects who reported using explicit strategies in the post-test questionnaire were eliminated from further analyses. Subjects classified as "test-aware" in the questionnaire produced the same results as those who were unaware of the connection between study and test items within each individual experiment (only one condition in Experiment 1 and the combined analysis across experiments revealed a statistically significant difference). (c) Subjects were also not informed of the test until after study and a cover story was used in the implicit test. Thus, incidental learning conditions were enforced in all three experiments. This was expected to minimize intentional retrieval attempts and avoid explicit access of the lexical form of the lure during rehearsal (see Seamon et al., 2002). (d) Test booklets included a large percentage of filler items (68% or 75%) to hide the nature of the test and ensure a low relatedness proportion. (e) Finally, subjects in the implicit test conditions were given a time constraint for completing trials to reduce the opportunities for explicit retrieval. Although all of the past studies that examined false priming in stem completion (McDermott, 1997; McKone & Murphy, 2000; Tajika et al., 2005) included controls for explicit contamination, it is possible that in these studies a small amount of explicit contamination might have occurred and thus allowed lure priming to be detected.

These issues could account for the discrepant results for stem completion found across the four studies (i.e., the current study, McDermott's, 1997, study, McKone & Murphy's, 2000, study, Tajika et al.'s, 2005, study). Small changes in the methodology could result in significant lure priming. This means that implicit priming effects for lures in perceptual tests with the DRM procedure may be more elusive than previously thought. Zeelenberg and Pecher's (2002) and Meade et al.'s (in preparation) results for a lexical decision task are consistent with this conclusion and the results of the current study. However, Tse and Neely (2005) did report significant priming for lures in a lexical decision task even when the critical lure was not tested immediately after a related list item. Some methodological differences between the present study and Tse and Neely's study might account for the discrepant results. First of all, in their study, intentional encoding conditions and relatively long exposure durations (i.e., 5 s per item) were in place, thus increasing the possibility that the critical lure was consciously accessed during rehearsal (e.g., Seamon et al., 2002). Second, subjects in Tse and Neely's series of experiments did repeated study-test trials, which might have increased their awareness of the structure of the DRM lists. Third, Tse and Neely's non-words (pseudo-homophones) may have altered the nature of their lexical decision task as inclusion of these non-words has been found to increase the likelihood of subjects accessing semantic level information and the lexical decision task can become a more conceptually driven task (see Zeelenberg & Pecher, 2002 for a similar argument). Finally, the delay between study and test in Tse and Neely's study was much shorter than in the current study, in which all lists were studied at once and a single test was administered at the end. From these results, it is unclear whether lexical forms of the lure items are available at study. This issue has implications for theories of false memory effects such as the activation-monitoring theory suggested by Roediger and colleagues (e.g., Gallo & Roediger, 2002; Roediger et al., 2001).

According to the activation-monitoring theory (Roediger et al., 2001), false memory effects in the DRM procedure are due to a combination of activation of the lures at study and source monitoring errors at test. Due to the semantic relatedness between list items and critical lure items, the list and lure items become "activated" in a similar fashion. This activation may be implicit or explicit and could include the lexical form of the lure items. At test, subjects make source-monitoring errors and mistakenly believe they have studied some of the lure items due to their high level of activation. Lexical activation of the critical lures would be consistent with the activation-monitoring explanation of false memory effects.

Recently, Lövdén and Johnasson (2003) examined false implicit memory in an anagram task using the DRM procedure. Their Experiment 1 results supported the "false" priming effects found in previous studies. In Experiment 2 they explored the cause of these effects by having subjects perform an articulatory suppression task during study. They found no lure priming when covert articulation was suppressed, indicating that verbal activation of the lures is necessary for false implicit memory. These results suggest specific constraints on the conditions necessary to produce lure priming and that activation of the lures at study is a likely explanation of false memory effects in the DRM paradigm. Based on Lövdén and Johansson's conclusion that articulatory access to the lures is important for perceptual "false" priming, one might expect that the graphemic response task used in the current study (with instructions to generate items that look or *sound* like the cue words) would be optimal for producing lure priming. However, no such priming was observed in the graphemic response task. One possibility is that articulatory access only occurs with intentional encoding conditions. Incidental encoding in the current study may have prevented this. Future studies will likely clarify the role of intentional or incidental learning in the DRM paradigm.

The current results are consistent with past findings (e.g., Meyer et al., 1974) that indicated that semantic priming is a short-term effect. Some recent studies (Becker, Moscovitch, Behrmann, & Joordens, 1997; Joordens & Becker, 1997) have reported long-term semantic priming, but there were some significant methodological differences between the present study and studies conducted by Becker et al. and Joordens and Becker. First of all, their studies used pure semantic associates (i.e., items selected from the same semantic category), whereas the DRM lists tend to combine semantically related and associatively related words (e.g., *bread* and *butter*). Second, the longest delay in Becker et al. and in Joordens and Becker was a lag of eight intervening items. Delays in the current study likely exceeded the time course of such effects as seen in past studies. Finally, the results in Joordens and Becker's study were not consistent across four experiments. In all, it appears that the evidence for long-term semantic priming is still inconsistent, and if semantic priming does occur, the effects are likely to be very small and only occur under specific circumstances.

In summary, the current study explored false implicit and explicit memory in stem completion and graphemic cued response tasks in the DRM paradigm. No significant priming was found for critical lures in either task, while false explicit memory was evident in both tasks. One criticism of the current study might deal with the relatively low rates of explicit false memory (which ranged from .03 to .27). As noted previously, these rates of false memory were significantly greater than the rate of intrusions for lures associated to non-studied lists (for semantically processed lists). Furthermore, several published studies have reported comparable rates of false recall. Robinson and Roediger (1997) reported rates of approximately 20% in two experiments following the study of nine-item lists. Hutchison and Balota (2005) reported rates of false recall between 5% and 10% for six-item lists and between 15% and 25% for 12-item lists. Thus, our rates of false cued recall are not inconsistent with those reported in other studies and, therefore, our failure to find priming for lures is probably not due to the fact that our materials did not elicit false memories for the critical lures in explicit tasks. In addition, these false memories were obtained under incidental learning conditions and after a fairly long delay.

In conclusion, our results are in direct contrast with results reported by McDermott (1997), McKone and Murphy (2000), and Tajika et al. (2005) for stem completion tasks and by Lövdén and Johnasson (2003) for an implicit anagram task (Experiment 1), but consistent with results of no lure priming in a lexical decision task reported by Zeelenberg and Pecher (2002), by McKone (2004) and by Meade et al. (in preparation). Due to small effect sizes for lure priming, false implicit memory may only be found in perceptual tests under certain conditions (e.g., with intentional encoding when subjects are more likely to be "test-aware") or may be due to explicit contamination. Alternatively, false implicit memory may be due to other factors related to awareness of the connection between study and test. Richardson-Klavehn and Gardiner (1996) suggest that conscious involuntary memory processes may occur when subjects are "test-aware", and these processes may have different effects on task performance than either voluntary conscious processes or automatic processes. As test awareness was not manipulated in the current study, it is difficult to draw specific conclusions as to the cause of awareness effects found in the current study. The exploration of the effects of "test-awareness" should be a focus of future studies in this area.

Acknowledgements

The authors wish to thank Jeffrey Wagman, Ewald Neumann, Chi-Shing Tse, and numerous anonymous reviewers, for helpful comments on earlier versions of this manuscript.

References

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1059–1082.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 657–668.

Coane, J. H., & McBride, D. M. (2002). Level of processing affects true and false memory similarly. *Presented at the meeting of the American Psychological Society*, New Orleans, LA.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Erdfelder, E., Faul, F., & Buchner, A. (1996). G*Power: A general power analysis program. *Behavior Research Methods, Instruments, and Computers, 28*, 1–11.

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

Gallo, D. A., & Roediger, H. L., III (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language, 47*, 469–497.

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.

Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory and Language, 52*, 1–28.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1083–1105.

Lövdén, M., & Johnasson, M. (2003). Are covert verbal responses mediating false implicit memory? *Psychonomic Bulletin and Review, 10*, 724–729.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological. 1: Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General, 123*, 107–128.

McDermott, K. B. (1997). Priming in perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin and Review, 4*, 582–586.

McKone, E. (2004). Distinguishing true from false memories via lexical decision as a perceptual implicit test. *Australian Journal of Psychology, 56*, 42–49.

McKone, E., & Murphy, B. (2000). Implicit false memory: Effects of modality and multiple study presentations on long-lived semantic priming. *Journal of Memory and Language, 43*, 89–109.

Meade, M. L., Watson, J. M., Balota, D. A., & Roediger, H. L. III. (in preparation). Spreading activation and false memories in the DRM paradigm: Retrieval mode is necessary.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory and Cognition, 2*, 309–321.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Available from http://www.usf.edu/FreeAssociation/.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin and Review, 7*, 158–162.

Richardson-Klavehn, A., & Gardiner, J. M. (1996). Cross-modality priming in stem completion reflects conscious memory, but not voluntary memory. *Psychonomic Bulletin and Review, 3*, 238–244.

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

Roediger, H. L., III, Balota, D. A., & Watson, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington, DC: American Psychological Association Press.

Roediger, H. L., III, & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam, The Netherlands: Elsevier.

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

Russell, W. A., & Jenkins, J. J. (1954). *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test.* (Tech. Rep. No. 11, Contract N8 ONR 66216, Office of Naval Research). Minneapolis: University of Minnesota.

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

Stadler, M. A., Roediger, H. L., III, & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory and Cognition, 27*, 494–500.

Tajika, H., Neumann, E., Hamajima, H., & Iwahara, A. (2005). Eliciting false memories on implicit and explicit memory tests after incidental learning. *Japanese Psychological Research, 47*, 31–39.

Tse, C., & Neely, J. H. (2005). Assessing activation without source monitoring in the DRM false memory paradigm. *Journal of Memory and Language, 53*, 532–550.

Zeelenberg, R., & Pecher, D. (2002). False memories and lexical decision: Even twelve primes do not cause long-term semantic priming. *Acta Psychologica, 109*, 269–284.