The roles of encoding and retrieval processes in associative and categorical memory illusions

Stephen A. Dewhurst, Emma Bould, Lauren M. Knott

Department of Psychology, Lancaster University, England

Craig Thorley

Department of Social and Psychological Sciences, Edge Hill University, England

Address for correspondence:

Dr Stephen A. Dewhurst

Department of Psychology

Lancaster University

Lancaster LA1 4YF

England

Email s.a.dewhurst@lancaster.ac.uk

Phone +44 1524 593835

Fax +44 1524 593744

Four experiments investigated the origin of associative and categorical memory illusions by comparing the effects of study and test associations on Deese/Roediger-McDermott (DRM) and categorized lists. Experiments 1 and 2 found that levels of false recognition with both list types were increased by manipulations that facilitated the generation of associates at study (blocked presentation of study lists and explicit instructions to generate associates of studied items). Experiments 3 and 4 showed that manipulations designed to increase test associations (test-induced priming and part-set cuing) did not increase levels of false memory with either list type. These findings indicate that false memories produced by both DRM and categorized lists are influenced by associations activated at study but not by associations activated at test.

Roediger and McDermott (1995) showed that powerful and compelling illusions of memory can be produced in a simple laboratory-based procedure based on lists of associated words. In what is now referred to as the DRM procedure (for Deese/Roediger-McDermott; Deese, 1959, Roediger & McDermott, 1995), participants study lists of words that are semantic associates of a nonpresented "critical lure". For example, participants study words such as *bed*, *rest*, *tired*, and *dream*, which are associates of the critical lure *sleep*. When asked to recall or recognize the studied words, participants frequently endorse the critical lures as old, with levels of false memory equalling or even exceeding levels of correct memory. Similar effects are observed in the category repetition procedure, in which participants study lists of words from taxonomic categories (e.g., animals, colours, parts of the body) from which some common exemplars are omitted (e.g., Dewhurst, 2001; Dewhurst, Barry, & Holmes, 2005; Seamon, Luo, Schlegel, Greene, & Goldenberg, 2000). Levels of false recall and false recognition produced by the category repetition procedure are typically smaller in magnitude than those produced by the DRM procedure (e.g., Park, Shobe, & Kihlstrom, 2005). Nevertheless, both procedures produce illusory memories that are often indistinguishable from true memories (e.g., Dewhurst & Farrand, 2004; Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997).

The DRM effect has been explained in terms of an activation-monitoring account (Roediger, Watson, McDermott, & Gallo, 2001), in which it is assumed that the critical lures are generated in response to the study lists.

This account is based on the “implicit associative responses” hypothesis proposed by Underwood (1965), according to which, participants spontaneously generate words that are semantic associates of the items presented to them. Underwood developed a continuous recognition procedure in which participants were presented with a sequence of words and had to decide whether a word appeared earlier in the sequence. He found that participants often judged a new word as old if it was semantically related to an earlier word. In the DRM procedure, participants who hear words such as *bed*, *dream*, and *wake* at study spontaneously generate the critical lure *sleep*. The DRM illusion occurs because participants are unable to remember the source (internally generated or externally presented) of the critical lures when they are presented at test (see Johnson, Hashtroudi, & Lindsay, 1993, for a review of the source monitoring framework). The activation-monitoring account can also explain the memory illusion produced by categorized lists (e.g., Dewhurst & Anderson, 1999).

Although the activation-monitoring account invokes both encoding and retrieval processes, previous research suggests that the associations that give rise to the memory illusion occur at study. For example, levels of false recall and false recognition are higher when DRM lists are presented in blocked rather than random sequences (McDermott, 1996; Toglia, Neuschatz, & Goodwin, 1999; Tussing & Greene, 1997) or in long rather than short lists (Robinson & Roediger, 1997). Encoding operations that encourage relational rather than item-specific processing (Hunt & Einstein, 1981) also increase false memories. For example, Arndt and Reder (2003) found that false recognition was higher when all items in a list were presented in the same font than when each word was presented in a unique font. Similarly, McCabe, Presmanes, Robertson, and Smith (2004) found greater levels of false recognition when participants were instructed to relate study items to one another than when they were instructed to focus on the unique characteristics of each word. These and other findings (see Gallo, 2006, for a review) show that encoding operations that encourage participants to make associations between study items enhance the false recall or false recognition of critical lures related to the list themes.

Previous studies have also considered the possibility that memory illusions are influenced by associations made during the test phase. For example, Roediger and McDermott (1995) examined output order in free recall and noted that critical lures were typically produced towards the end of the recalled words, suggesting that they may have been cued by the words that were correctly recalled. They also noted that critical lures in recognition tests may have been primed by studied items presented earlier in the list. However, Roediger et al. (2001) found a negative correlation between correct and false recall and concluded that test associations play little role in the DRM illusion.

Other studies have attempted to influence false recall and false recognition by inducing associations at test. For example, Reysen and Nairne (2002) used the part-set cuing procedure in which a subset of studied items was presented as a cue to recall the remaining items. Part-set cuing has previously been found to reduce correct recall (see Nickerson, 1984, for a review). As Reysen and Nairne observed, this is a counterintuitive finding as one would expect list cues to facilitate recall. If the DRM effect is influenced by associations made at test, then one would also expect false recall to be increased by part-set cues. However, Reysen and Nairne found that false recall, as well as correct recall, was reduced by test cues. Findings consistent with these were also reported by Kimball and Bjork (2002) and Kimball, Bjork, Bjork, and Smith (2008).

Some evidence for a role of test associations in the DRM effect comes from studies that used a test-induced priming procedure (e.g., Coane & McBride, 2006; Dodd, Sheard, & MacLeod, 2006; Marsh & Dolan, 2007; Marsh, McDermott, & Roediger, 2004). This procedure attempts to induce false memories at test by manipulating the number of studied items (or unstudied but related items) that precede the critical lure in the recognition test. Both Dodd et al. and Marsh et al. found that test primes did not increase levels of false recognition for studied lists, though Marsh et al. found that levels of false recognition increased above baseline for nonstudied lists. In contrast, Marsh and Dolan (2007) found an increase in false recognition when participants had to make old/new decisions before a 750 msec response deadline (but see Dodd et al. for a null effect under speeded response conditions). Marsh and Dolan suggested that test primes contribute to the DRM illusion, but only when monitoring strategies are impaired. Coane and McBride found an increase in false recognition under self-paced test conditions when critical lures were preceded by 6 or 12 studied items. However, they acknowledged that the effects of activation at test are weaker than the effects of activation at study and that the DRM effect is driven largely by processes that occur at study.

There is also evidence that false memories produced by the category repetition procedure, like those produced by the DRM procedure, are influenced by processes that occur at study. For example, dividing attention at study reduces the false recognition of critical lures with both categorized and DRM lists (Dewhurst et al., 2005; Dewhurst, Barry, Holmes, Swannell, & Bathurst, 2007; Knott & Dewhurst, 2007). Dewhurst et al. (2005) also found that the false recognition of category lures increased when participants were explicitly instructed to make associations to study lists. In addition, increasing the number of category exemplars presented at study increases false recognition rates (Dewhurst & Anderson, 1999). These findings suggest that categorical memory illusions, as well as associative memory illusions, are influenced by associations activated at study.

The effects of encoding operations on categorical memory illusions are, however, at odds with the suggestion of Smith, Gerkens, Pierce, and Choi (2002). They investigated whether associative and categorical memory illusions are the product of associations activated at the study phase, which they termed the Kirkpatrick hypothesis (after Kirkpatrick, 1894), or of associations activated at test, which they termed the Deese hypothesis (after Deese, 1959). In their first experiment, Smith et al. found that critical lures from DRM lists were ten times more likely than critical lures from categorized lists to be produced in a word association task. In their second and third experiments, they used an indirect priming paradigm in which participants studied DRM and categorized lists and were then given a stem completion task. Smith et al. found large priming effects with DRM lists, whereby participants frequently completed the word stems with critical lures. In contrast, no priming effects were observed with categorized lists, relative to a baseline condition. However, the DRM and categorized lists produced equivalent levels of false recall. In their fourth experiment, they found a priming effect in categorized lists when participants were instructed to complete the stems with words from the study lists, but not when instructed to complete them with the first word that came to mind.

Smith et al. (2002) concluded from these findings that false memories in the DRM procedure are influenced by associations made during the study phase, whereas false memories in the category repetition procedure can be produced by semantic processes that occur at test. They suggested that the difference between DRM and categorized lists may be due to the greater backwards associative strength (BAS) of DRM lists relative to categorized lists. BAS refers to the tendency of the critical lures to be generated in response to list items, and has been found to be the main predictor of false recall and recognition in the DRM procedure (see Roediger et al., 2001).

The aim of the present study was to investigate the roles of study and test processes on associative and categorical memory illusions. We report a series of experiments in which we manipulated participants' opportunity to make associations either at study or at test. If Smith et al.'s (2002) suggestion is correct, conditions that facilitate associations at study should exert a greater influence on DRM than on categorized lists, whereas conditions that facilitate associations at test should exert a greater influence on categorized than on DRM lists. Alternatively, if false memories produced by both list types are influenced by study associations but not by test associations, then both should be enhanced by experimental manipulations that facilitate associations at study but not by manipulations that facilitate associations at test.

Experiment 1

Experiment 1 investigated the role of study associations by manipulating the organization of the study lists. Tussing and Greene (1997) found that levels of false recognition were higher when the lists were presented in blocked rather than random sequences. In contrast, Dewhurst and Anderson (1999) found no significant difference between blocked and random presentations in false recognition using categorized lists. These findings appear to support the views of Smith et al. (2002) regarding the origins of the DRM and category repetition effects. However, the mean false alarm rates presented by Dewhurst and Anderson show that false recognition rates were numerically higher following blocked rather than random study presentation, particularly when lists of eight exemplars were studied (list length was manipulated within-groups such that participants saw one, four, or eight items per category). It is possible that presentation mode (blocked versus random) will influence the false recognition of category exemplars when longer lists are studied. We investigated this possibility by comparing the effects of presentation mode on DRM and categorized lists of 10 items each.

In Experiments 1 to 3, the primary analyses were conducted on the overall proportions of hits and false alarms. Following Roediger and McDermott (1995: Experiment 2), we also investigated the subjective experience of false recognition by asking participants to categorize their positive decisions as either remember or know responses (Gardiner, 1988; Tulving, 1985). Some previous studies have shown selective effects of manipulated variables on either false remember responses (e.g., Dewhurst et al., 2001) or false know responses (e.g., Seamon, Luo, and Gallo, 1998)., which suggests that the remember-know procedure can provide a finer-grained analyses than standard old/new recognition tests. It is important to note, however, that we used the remember-know procedure simply to determine the subjective experience of recognition and do not view it as a pure measure of any underlying systems or processes (see Gardiner & Richardson-Klavehn, 2001, for further discussion of this issue). Guess responses were not analyzed as they are typically made at or below chance levels (Gardiner, Ramponi, & Richardson-Klavehn, 2002).

Method

Participants. Participants were 80 (43 female, 37 male) undergraduate students from Lancaster University, aged 18-30. All participants in this and the following experiments spoke English as their first language.

Stimuli. The stimuli consisted of 20 DRM lists and 20 categorized lists of 10 words each. DRM lists were taken from the norms produced by Stadler, Roediger, and McDermott (1999) and categorized lists were taken from Van Overschelde, Rawson, and Dunlosky (2004). High frequency exemplars were used as critical lures for the categorized lists. Mean BAS values taken from the University of South Florida free associations norms (Nelson, McEvoy, & Schreiber, 1998) were .17 for the DRM lists and .03 for the categorized lists. The recognition test contained 100 items; 40 studied items (two from each list), the 20 critical lures, and 40 unrelated lures.

Design. List type and presentation format were manipulated between-groups such that half the participants studied 20 DRM lists and the other half studied 20 categorized lists. Within each group of participants, half studied sequences of lists presented in blocked format and the other half studied sequences in which items from the different lists were presented in a random order. In the random condition, words from the same list were separated by at least two words from other lists. The dependent variables were the numbers of correct and false remember and know responses.

Procedure. Participants were tested at individual workstations in groups of up to eight people. The experiment was presented on Apple Macintosh computers using custom written software. Participants were told that the 200 words would be presented in lists and that there were 20 lists, with 10 words in each list, and that at the end of each list the message “next list” would be shown for two seconds. Participants were instructed that they should try to learn the words in readiness for a memory test. A 5-minute filled retention interval followed, in which participants solved multiplication problems. They were then given the instructions for the recognition test. Test items appeared one at a time on the computer, and participants pressed response keys labelled O and N for old and new, followed by keys marked R, K and G to indicate remember, know, and guess responses. Instructions for remember, know, and guess responses were taken from Dewhurst and Anderson (1998). Briefly, participants were instructed to make a remember response if they recollected some aspect of the item's study presentation, such as an image or thought they experienced at the time; a know response if the word felt familiar but they were unable to recollect any detail of its study presentation; or a guess response if they were unsure whether or not the word had appeared at study. Test items were presented in a different random order for each participant.

Results and Discussion

Table 1 shows mean proportions of "old" responses to targets, critical lures, and unrelated lures as a function of list type and presentation format, plus correct and false remember, know, and guess responses. Proportions of overall hits and false alarms were analysed in separate 2 (list type: DRM vs categorized) x 2 (order: blocked vs random) between-groups ANOVAs. Correct recognition was higher following blocked than random presentation, *F* (1,76) = 12.92, *MSE* = .02, partial η^2 = .15, and higher for categorized than DRM lists, *F* (1,76) = 8.23, *MSE* = .03, partial η^2 = .10. The interaction was not significant, $F < 1.4$. Similar analyses were conducted on correct remember and know responses. Correct remember responses showed a significant main effect of presentation order, F (1,76) = 9.53, *MSE* = .03; η_p^2 = .11, and were higher with blocked than with massed presentations. The main effect of list type was not significant, F (1,76) = 2.92, *MSE* = .03, p = .09, η_p^2 = .04, nor was the interaction between list type and order, $F < 1.20$. Correct know responses showed nonsignificant effects of list type, F (1,76) = 2.31, *MSE* = .01, p = .13, η_p^2 = .03, and order, $F < 1$, and a nonsignificant interaction, $F < 1$.

Please insert Table 1 about here

Our main interest was in the effects of presentation order on the false recognition of critical lures for DRM and categorized lists. Overall false recognition rates were higher following blocked than random presentation, F (1,76) = 11.89, *MSE* = .04, η_p^2 = .14, and higher for DRM than for categorized lists, F (1,76) = 18.14, *MSE* = .04, η_p^2 = .19. The interaction was not significant, $F < 1$. As our aim was to determine whether presentation order influenced both DRM and categorized lists, we performed separate independent-samples t-tests for each list type. These confirmed that false remember responses were reliably higher following blocked than random presentation for both DRM lists, t (38) = 2.22, and categorized lists, t (38) = 2.67.

These effects were mirrored in false remember responses, which showed significant main effects of list type, *F* (1,76) = 8.98, *MSE* = .02, η_p^2 = .11, and order, *F* (1,76) = 20.56, *MSE* = .02, η_p^2 = .21. False know responses were higher for DRM than for categorized lists, *F* (1,76) = 4.64, *MSE* = .02, η_p^2 = .06, but were not reliably affected by presentation order, *F* < 1. The only significant finding in the unrelated distractors was a significant main effect of list type in false remember responses, *F* (1,76) = 4.84, *MSE* = .01, η_p^2 = .06, which showed that false remember responses to unrelated lures were significantly higher in the groups who studied DRM lists.

The findings from Experiment 1 replicate the finding by Tussing and Greene (1997) that false recognition in the DRM procedure is greater when study lists are presented in blocked rather than random sequences. The crucial finding from Experiment 1, however, was that blocked study lists also increased false remember responses for categorized lists. The parallel effects of blocked presentation on DRM and categorized lists suggest that false memories in both procedures are enhanced by associative processes that occur at study. The finding that DRM lists produce higher levels of false recognition than categorized lists is most likely due to differences in mean BAS. This will be considered further in the General Discussion.

Experiment 2

Dewhurst et al. (2005) showed that false remember responses in the category repetition procedure increased when participants were explicitly instructed to generate associates of the words presented at study. Participants were shown a series of category labels, each followed by 8 exemplars, and instructed to generate other members of the target categories during presentation. Participants showed increased levels of false remember responses to non-presented exemplars, relative to a second group of participants who were instructed simply to read the words. This finding provides strong support for the view that false memories in the category repetition procedure are influenced by the activation of associates at encoding. The aim of Experiment 2 was to compare the effect of instructions to make associations on DRM and categorized lists. Half the participants studied categorised lists and the remaining half studied DRM lists. Within each group of participants, half were instructed to generate associates of the items presented during the study phase and half were instructed simply to read the words silently in preparation for a memory test.

Method

Participants. Participants were 120 (76 female, 44 male) undergraduate and postgraduate students from Lancaster University, aged 18-36.

Stimuli. Stimuli consisted of the 20 DRM and 20 categorized lists used in Experiment 1. Sixty participants studied categorised lists and the remaining 60 participants studied DRM lists. The lists were presented in a blocked sequence on a Microsoft PowerPoint presentation. Each word remained on the screen for 1 second, with an inter-stimulus interval of 1 second. At the end of each block of words, a screen with the words ‘next list’ was displayed for 3 seconds. The recognition tests consisted of 40 target items, 20 related lures, and 20 unrelated lures presented on two sides of a response sheet in random order across two columns. The letters *R*, *K*, and *G* appeared to the right of each test item to indicate remember, know, and guess.

Design. List type (DRM or category) and instruction (read or generate) were both manipulated between-groups. The dependent variables were the number of correct and false remember and know responses.

Procedure. Participants were told that they would be shown a list of words for which they would later be given a recognition test. Within each list condition (DRM or categorized), 30 participants were instructed to read the words silently in anticipation for a memory test (read group) while the remaining 30 participants were instructed to generate associates of the study items as they were presented (generate group). After the presentation of the 20 lists there was a 5-minute filled retention interval after which participants were given the instructions for the recognition test.

Results and Discussion

Table 2 shows mean proportions of “old” responses to targets, critical lures, and unrelated lures as a function of list type and encoding instructions, plus remember, know, and guess responses. Targets, critical lures, and unrelated distractors were analyzed in separate 2 (list type: DRM or categorized) x 2 (instructions: read or generate) between-groups ANOVAs. The analysis of total hits showed nonsignificant main effects of list type and instructions and a nonsignificant interaction, all *F* < 1.2. The only significant effect in the analyses of remember and know responses was a main effect of instruction on know responses, *F* (1,116) = 4.39, *MSE* = .01, η_p^2 = .04, which were greater following read rather than generate instructions.

Please insert Table 2 about here

The analysis of critical lures showed significant main effects of list type, *F* (1,116) = 45.97, *MSE* = .04, η_p^2 = .29, and instructions, *F* (1,116) = 4.00, *MSE* = .04, η_p^2 = .03. The interaction did not reach significance levels, *F* < 1. The main effect of instruction was mirrored in false remember responses, *F* (1,116) = 10.12, *MSE* = .03, η_p^2 = .08, which were increased by instructions to generate. False remember responses were also greater for DRM than for categorized lists, *F* (1,116) = 50.54, *MSE* = .03, η_p^2 = .30. Again, the interaction between list type and instruction was not significant, *F* < 1, indicating that false remember responses for both list types were enhanced by explicit instructions to generate. Separate t-tests confirmed that instructions to generate significantly increased false remember responses for both DRM lists, *t* (58) = 2.35, and categorized lists, *t* (58) = 2.38. False know responses were not reliably affected by list type, *F* (1,116) = 2.99, *MSE* = .02, *p* = .09, η_p^2 =.03, or by instruction, *F* < 1, nor was the interaction significant, *F* < 1. False recognition of unrelated lures was low and not significantly affected by list type or instructions to generate.

The main finding from Experiment 2 was that false remember responses were enhanced by explicit instructions to generate associates at study. These findings replicate those reported by Dewhurst et al. (2005) with categorized lists and extend the effect to DRM lists. The finding that instructions to generate influenced only remember responses is consistent with the findings reported by Dewhurst et al. They suggested that explicit instructions influence the likelihood that generated associates reach conscious awareness, rather than the generation process itself, thereby influencing false remember responses (which require conscious awareness of generated items) but not false know responses (which do not require conscious awareness of generated items). This interpretation is consistent with findings that false remember responses are reduced under conditions that restrict the conscious generation of associates, while false know responses are reliably observed following the nonconscious activation of associates (Seamon, Luo, & Gallo, 1998).

Considered together, the results of Experiments 1 and 2 provide converging evidence that false recognition in both the DRM and category repetition procedures is influenced by associations activated at study. In Experiments 3 and 4 we investigated whether false memories in the DRM and category repetition procedures are influenced by associations activated at test.

Experiment 3

Experiment 3 investigated the effects of test-induced priming on false recognition for DRM and categorized lists. As discussed in the Introduction, there is evidence from previous studies that the DRM effect is enhanced by test primes, particularly when testing occurs under conditions that inhibit monitoring strategies, such as requiring participants to respond before a 750 msec deadline. The aim of Experiment 3 was to compare the effects of test-induced priming on false recognition produced by DRM and categorized lists. If Smith et al. (2002) are correct in their suggestion that false memories in the category repetition procedure are influenced by semantic processes that occur at retrieval, then levels of false recognition should increase with the number of category exemplars that precede the critical lure at test.

Following Marsh and Dolan (2007), we used two test conditions; one group of participants completed the recognition task at their own pace while a second group were tested under speeded response conditions. In Marsh and Dolan's study, participants made old/new decisions at test. In the present study we again used the remember-know procedure to measure the subjective experience of recognition. As Gardiner, Ramponi, and Richardson-Klavehen (1999) found that participants can reliably make both remember and know responses under speeded response conditions (see also Gardiner, Konstantinou, Karayianni, & Gregg, 2005) we anticipated no difficulties in obtaining remember and know responses in Experiment 3. Our use of the remember-know procedure also allowed us to determine whether test induced primes induce an illusion of recollection or enhance the subjective familiarity of critical lures.

Method

Participants. Participants were 120 (87 female, 33 male) undergraduate and postgraduate students from Lancaster University, aged 18-27.

Stimuli. The stimuli consisted of 36 DRM and 36 categorized lists of 10 items each. Mean BAS values (from Nelson et al., 1998) were .17 for the DRM lists and .05 for the categorized lists. Each list was divided into two sets of 18 (Set 1 and Set 2) and was presented to half the participants in each group. The recognition test contained 297 items; 6 list items from each of the 36 lists (18 studied and 18 non-studied), the 36 critical lures (18 studied and 18 non-studied) and 45 unrelated lures.

Design. A mixed design was used in which list type (DRM versus categorized) was manipulated between-groups and the number of test primes (0, 3 and 6) was manipulated within-groups. Half the participants studied the categorized lists and were shown either category set 1 or category set 2, and the other half studied the DRM lists and were shown either DRM set 1 or DRM set 2. Thus each participant studied 18 lists of words.

The critical manipulation was the number of studied items that preceded the critical lure in the recognition test (zero, three, or six). Counterbalancing was achieved by rotating each critical lure (studied and non-studied) through three positions, and in each of the three positions one third of the critical lures were shown before the six studied items, one third after three of the studied items, and the other third after all six of the studied items. The dependent measures were the numbers of correct and false remember and know responses.

Procedure. The experiment was presented on Apple Macintosh computers using custom written software. Participants were informed that they would be shown 18 lists of 10 words each, for which they would later be given a memory test. Each word was shown for one second followed by a blank screen for one second, and at the end of each list of words a screen displaying “next list” was shown for two seconds. Lists were presented in blocked order for all participants. At the end of the study phase there was a 5-minute filled interval, after which participants were given the instructions for the recognition test.

Results and Discussion

The proportions of total hits and false alarms plus remember, know, and guess responses are shown in Table 3. Hits were analyzed in separate 2 (list type: DRM vs. categorized) X 2 (response deadline: self-paced vs. speeded) between-groups ANOVAs. A significant main effect of response deadline was observed, F (1,116) = 48.94 MSE = .03, η_p^2 = .30, whereby correct recognition was higher in the self-paced than in the speeded condition. Neither the main effect of list type, $F < 1.5$, not the interaction with response deadline, $F < 1$, were significant. The effect of response deadline was also observed in separate analyses of remember responses, F (1,116) = 28.33, MSE = .78, η_p^2 = .20, and know responses, F (1,116) = 4.47, MSE = .01, η_p^2 = .04. A similar analysis of unrelated lured showed that they were marginally higher for DRM than for categorized lists, F (1,116) = 2.88, *MSE* = .02, p = .09, η_p^2 = .02, but were not reliably affected by response deadline, $F < 1.6$.

Please insert Table 3 about here

Our main focus was on the effect of test position on the false recognition of critical lures. Preliminary analyses indicated that false recognition scores were higher for studied than for unstudied lists, F (1,116) = 153.32, *MSE* = .04, η_p^2 =.57, higher in the self-paced than in the speeded response condition, F (1,116) = 7.28, *MSE* = .08, η_p^2 = .06, and higher for DRM than for categorized lists, F (1,116) = 12.89, *MSE* = .08, η_p^2 = .10. In order to explore the effects of test primes, the data were analysed in separate 2 (list status: studied vs. nonstudied) x 3 (number of test primes: 0 vs. 3 vs. 6) repeated measures ANOVAs on false recognition scores for DRM and categorized lists in the self-paced and speeded response conditions. As false recognition rates were always significantly higher for studied than for unstudied lists, only the effects of number of test primes and interactions with list status are reported.

For DRM lists in the self-paced condition, there was a significant effect of test primes, F (1,29) = 3.87, *MSE* = .04, η_p^2 = .12. Pairwise comparisons showed that false recognition was significantly reduced when critical lures were preceded by 6 primes rather than by 3 primes. A significant main effect of test primes was also observed in the speeded DRM condition, F (1,29) = 3.91, *MSE* = .03, η_p^2 = .12, and pairwise comparisons showed that false recognition was reduced when critical lures were preceded by 6 primes than by 3 or 0 primes. The interaction between list status and number of primes was not significant, $F < 1$. For categorized lists in the self-paced condition, neither the main effect nor the interaction were significant, both $F < 1$. Similarly in the speeded categorized condition, neither the main effect of primes $F\ (1,29) = 2.15$, $MSE = .04$, $p = .13$, $\eta_p^2 = .07$ nor the interaction, $F < 1.3$, was significant.

Separate analyses were again conducted on remember and know responses. For DRM lists in the self-paced condition, the analysis of false remember responses showed a significant main effect of test primes, $F\ (2,58) = 3.31$, $MSE = .03$, $\eta_p^2 = .10$, and pairwise comparisons showed that false remember responses were reliably lower after 6 than after 3 test primes, $p < .05$. The interaction with list status was not significant, $F\ (2,58) = 2.40$, $MSE = .02$, $\eta_p^2 = .08$. For DRM lists in the speeded response condition, false remember responses showed no reliable effect of test primes, $F < 1.8$, but a marginally significant interaction with list status, $F\ (2,58) = 3.11$, $MSE = .02$, $p = .05$, $\eta_p^2 = .10$. Pairwise comparisons showed that false remember responses to studied lists were significantly reduced by 6 relative to 3 primes, $p < .05$, and marginally reduced by 6 relative to 0 primes, $p = .07$. The analyses of categorized lists showed no reliable effects of test primes in either self-paced, $F < 2$, or speeded response conditions. Although it may appear from Table 3 that false remember responses for studied lists increased with the number of primes in the self-paced condition, a one-way ANOVA indicated that the effect was not significant, $F < 1.2$, and nor were any of the pairwise comparisons.

The only significant effect to emerge from the analyses of know responses to DRM lists was a significant status x test primes interaction in the self-paced condition, F (2,58) = 4.52, MSE = .02, η_p^2 = .14. Pairwise comparisons showed a marginally significant decrease in false know responses to lures from nonstudied lists after 6 primes compared to 3, $p = .09$, but no reliable effects of test primes for studied lists. For categorized lists, know responses in the self-paced condition were significantly lower after 6 primes than after 0, $p < .05$, and 3, $p = .06$, primes.

The findings from Experiment 3 provide no evidence that false recognition is influenced by associations made at test. In contrast to the findings of Marsh and Dolan (2007), test primes reduced false recognition under speeded response conditions. The findings of Experiment 3 are also inconsistent with those of Coane and McBride (2006), who found increased levels of false recognition following 6 or 12 primes under self-paced conditions. One possible explanation for the discrepancy between our findings and theirs is our use of the remember-know procedure. Requiring participants to consider the experiential basis of their recognition decisions is likely to have heightened their memory monitoring strategies, even under speeded response conditions. This is supported by the observation that overall levels of false recognition in Experiment 3 were lower than those reported by Coane and McBride and by Marsh and Dolan. Mean recognition scores reported by Marsh and Dolan (who used the same numbers of test primes as the present study) for studied and unstudied lists in the speeded condition were .60 and .35. In comparison, the corresponding recognition scores in the present study were .39 and .20. Marsh and Dolan suggested that test cues increase false recognition under conditions that impair monitoring strategies. The findings from Experiment 3 are consistent with this view in showing that the effect is eliminated when participants are encouraged to monitor their memory performance. In contrast to the DRM lists, the analysis of categorized lists showed that test primes reduced false know responses, though only under self-paced conditions.

Experiment 4

Experiment 4 investigated the effects of part-set cuing on false recall with DRM and categorized lists. Reysen & Nairne (2002) found that part-set cues reduced false recall (but see Marsh et al., 2004, who found no effect of part-set cues in false recall). In their second experiment, Reysen and Nairne tested the possibility that the part-set effect occurs because the retrieval cues disrupt participants' preferred retrieval strategies (see Basden & Basden, 1995). They did this by presenting cues that were either consistent with participants' retrieval strategies (even numbered list items presented in the same order as at study) or inconsistent with them (random list items). They found that the consistency of the cues influenced correct recall (greater disruption with inconsistent cues) but had no effect on the false recall of critical lures. Reysen and Nairne suggested that the reduction in false recall following the presentation of part-cues is not due to a disruption of preferred retrieval strategies but to an enhancement of source monitoring. Specifically, the retrieval cues reinstate the encoding context, thereby increasing participants' ability to determine whether the test items had been externally presented or internally generated.

Experiment 4 followed the procedure described by Reysen and Nairne (2002) and compared the effects of recall cues on DRM and categorized lists. If false recall in the category repetition procedure is influenced by semantic processes that occur at test, then levels of false recall should be enhanced by part-set cues. Participants were presented with eighteen 12-item DRM or categorized lists. After studying each list, the participants were instructed to recall the words without any cues or with six cues taken randomly from the list. To our knowledge, no previous studies have investigated the effects of part-set cuing on false recall using categorized lists. Presenting category labels or exemplars from studied lists has been shown to produce part-set cuing effects in correct recall (e.g., Marsh, Dolan, Balota, & Roediger, 2004; Roediger, 1978; Slamecka, 1968), but no effects in false recall have been reported.

Method

Participants. Participants were 56 (38 female, 18 male) undergraduate students from Lancaster University, aged 18-30.

Design and Stimuli. The stimuli consisted of 18 DRM lists and 18 categorized lists, each containing 12 words. Half the participants studied categorized lists and half studied DRM lists. Mean BAS levels (from Nelson et al., 1998) were .20 for DRM lists and .04 for categorized lists. Following the presentation of each list, participants solved simple multiplication problems for 15 seconds before completing a free uncued recall task or a cued recall task. Across the course of the experiment, each participant received 9 uncued recall tests (requiring free recall of the just-studied list) and 9 cued recall tasks (in which six list items were visually displayed on a computer screen and they were asked to recall the remaining words from the list). Participants wrote the answers in a booklet. The assignment of each list to the two cue conditions was counterbalanced, so that all lists appeared equally often in each test condition. The order of presentation for each list and cue type was determined randomly for each participant. The six cues were chosen randomly from each list and were presented in a random order.

Procedure. Participants were presented with 18 lists of words, each consisting of 12 words presented visually at a rate of 1.5 seconds per word, with a 1 second gap between each presentation. After the presentation of lists, participants were instructed to answer the multiplication problems, and after 15 seconds recall the just-studied list. For the cued recall task, six cue items were displayed in a column on the computer screen, and the participants were asked to write in their booklet as many of the remaining list items as possible. For the uncued recall task, participants were prompted to write down as many words from the list as they could remember. Both the cued and uncued recall tasks were self-paced. Participants were instructed to press the space bar when they could recall no further items from the list. The next list was then displayed and the sequence was repeated.

Results and Discussion

The mean proportions of correct and false recall as a function of cue condition and list type are shown in Table 4. Correct and false recall scores were analyzed in separate 2 (list type: DRM vs. categorized) X 2 (cue condition: uncued vs. cued) mixed model ANOVAs with repeated measures on the second factor. There was a significant main effect of cue condition on correct recall, F (1, 54) = 46.59, MSE = .01 η_p^2 = .46, indicating that recall was greater on the uncued tests than on the cued tests. The main effect of list type was also significant, F (1, 54) = 9.89, MSE = .03 η_p^2 = .16, with a higher proportion of correct recall for categorised lists relative to DRM lists. The interaction was not significant, $F < 1$.

The analysis of false recall showed a significant main effect of list type, F (1, 54) = 22.19, MSE = .07 η_p^2 = .29, with a higher incidence of false recall for DRM than for categorized lists. There was no effect of cue condition, $F < 2$, but the list type by cue condition interaction was significant, F (1, 54) = 6.35, MSE = .02 η_p^2 = .11. Analysis of simple main effects indicated that for DRM lists there was a higher incidence of false recall on uncued tests than on cued tests, F (1, 54) = 6.65, MSE = .02 η_p^2 = .20, however there was no significant effect of cue condition for categorised lists F

< 1.

The results of Experiment 4 show that the false recall of DRM critical lures is reduced when studied items are presented as recall cues. This is consistent with the findings of Reysen and Nairne (2002). In contrast, no reliable effect of test cues was observed with categorized lists. Reysen and Nairne suggested that part-set cues reinstate the encoding context and thereby reduce false recall by enhancing participants' monitoring strategies. The finding from Experiment 4 that part-set cuing did not reduce false recall with categorized lists is inconsistent with this account, as participants are also required to determine the source of test items after studying categorized lists. However, it is possible that participants' monitoring strategies are influenced by list type. In categorized lists, study items are typically nouns whose membership of the target category is obvious. In contrast, DRM lists consist of a range of word types and often feature more than one theme (see Howe, Wimmer, & Blease, in press). Participants may be able to reduce false recall with categorized lists by using a recall-to-reject strategy, whereby they are able to recall the studied items and thus reject the critical lures (e.g., Brainerd, Reyna, Wright, & Mojardin, 2003; Gallo, 2004; Tulving, 1983). In contrast, participants may be less able to use this strategy with DRM lists because the multiple themes within the lists make it harder to recall the studied words. It is possible that the enhancement of monitoring strategies by part-set cues is effective only when test conditions do not facilitate a recall-to-reject strategy, as is the case with DRM lists.

General Discussion

The aim of the present study was to investigate the respective roles of encoding and retrieval processes in associative and categorical memory illusions. Previous research has shown that false memories produced by the DRM procedure are influenced by associations activated at study (Roediger & McDermott, 1995, Roediger et al., 2001) and are only weakly influenced by processes occurring at test (Coane & McBride, 2006; Marsh & Dolan, 2007). Smith et al. (2002) suggested that false memories produced by the category repetition procedure differ from those produced by the DRM procedure in that they are more likely to be elicited by semantic processes occurring at test. In contrast to this view, the present findings suggest that false memories in both procedures are influenced by associations activated at study but not by associations activated at test.

Experiments 1 and 2 employed encoding manipulations that have previously been found to increase false memories. In Experiment 1, word lists were presented either blocked or in a random sequence. Consistent with previous findings (McDermott, 1996; Toglia et al., 1999; Tussing & Greene, 1997), levels of false memory were higher for participants who studied blocked lists. In Experiment 2 word lists were presented in either a control condition in which participants read the words silently or a generate condition in which they were explicitly instructed to make semantic associations to the words as they appeared on the screen. Consistent with the findings of Dewhurst et al. (2005), false recognition of critical lures was enhanced by instructions to generate. Crucially for the argument we are presenting, the manipulations of list structure and generation produced the same patterns of results in DRM and categorized lists, suggesting that false memories with both list types are influenced by associations activated at study.

Experiments 3 and 4 employed retrieval manipulations designed to increase the activation of critical lures at test. In Experiment 3, critical lures in the recognition test were preceded by zero, three, or six studied items from the corresponding list. Previous research has shown that levels of false recognition can be enhanced by test-induced primes under both self-paced recognition conditions (Coane & McBride, 2006) and speeded response conditions (Marsh & Dolan, 2007). In contrast, we found that test primes did not increase false recognition under self-paced conditions and reduced false recognition under speeded conditions. This pattern bears out the observation by Tse and Neely (2007) that studies of within-test priming have produced inconsistent results, with some studies reporting null effects, some reporting reductions in false alarms, and others reporting increases in false alarms. Marsh and Dolan suggested that test associations can contribute to associative memory illusions under conditions that impair memory monitoring. If so, the findings of Experiment 3 indicate that their influence is relatively weak and easily eliminated.

Experiment 4 used a part-set cuing paradigm in which participants were asked to recall each list and were prompted either by zero or six studied items. As in Experiment 3, the crucial finding was that part-set cuing did not increase false memory with either DRM or categorized lists, and in fact reduced false recall with DRM lists. Overall, the findings from the present study are consistent with the activation-monitoring account proposed by Roediger et al. (2001). According to this account, it is the associations activated at study that give rise to the DRM illusion. The parallel effects observed with DRM and categorized lists in Experiments 1 and 2 indicate that the activation-monitoring account can explain both associative and categorical memory illusions.

A consistent finding across all experiments was that levels of false memory were significantly higher for DRM than for categorized lists. These differences occurred despite equivalent levels of correct recognition in Experiments 2 and 3 and higher levels of correct recognition and recall for categorized lists in Experiments 1 and 4. The lower levels of false memory for categorized lists are likely due to the greater BAS of DRM lists, as the high numbers of lists required in these experiments made it impossible to match them for BAS. Yet despite the higher levels of false recognition for the DRM lists, both list types showed the same overall pattern of effects in that false recognition rates were enhanced by encoding conditions that facilitated the generation of associations at study but were not enhanced by priming at test. In spite of differences in the strength of the associations that underlie associative and categorical memory illusions, the findings from the present study suggest that they are both influenced by associations made at study but not by associations activated at test.

This conclusion is inconsistent with the suggestion by Smith et al. (2002) that categorical memory illusions can be elicited at test rather than at study. Smith et al. suggested that false memories in the DRM procedure are influenced by associations made during the study phase, whereas false memories in the category repetition procedure can be produced by semantic processes that occur at test. They based this suggestion on their finding that critical lures of DRM lists were more likely to be generated in word association and stem completion tasks than critical lures of categorized lists, even though levels of false recall were equivalent. Smith et al. suggested that the differences may be due to the greater BAS of DRM lists relative to categorized lists. The findings of Smith et al. suggest potentially important differences between DRM and categorized lists that warrant further investigation. However, the findings from the present study indicate that the memory illusions produced by both DRM and categorized lists are influenced by associations activated at study but not by associations activated at test.

We are not claiming that retrieval processes play no role in associative and categorical memory illusions. The activation-monitoring account (Roediger et al., 2001) invokes a combination of encoding and retrieval processes, whereby participants make source monitoring errors at test and are unable to determine whether critical lures were externally presented or internally generated at study. Reysen and Nairne (2002) suggested that part-set cues reduce false recall by enhancing participants’ monitoring strategies. The results of Experiment 3 also indicate that the increase in false recognition by test primes reported by Coane and McBride (2006) and by Marsh and Dolan (2007) is eliminated when participants are required to make remember/know decisions at test. As discussed above, our use of the remember-know procedure is likely to have enhanced participants' monitoring strategies, even in the speeded response condition, thereby eliminating the effects reported by Coane and McBride and by Marsh and Dolan. Considered together, these findings suggest that experimental manipulations of test processes influence participants' source monitoring strategies rather than the activation of associates.

The role of source monitoring failures in false recognition was demonstrated in a recent study by Knott and Dewhurst (2007), who found that false remember responses increased when participants were instructed to perform random number generation during the recognition test. Knott and Dewhurst suggested that dividing attention at test impaired the controlled source monitoring strategies that would otherwise have prevented false recognition, forcing participants to rely on less stringent automatic source monitoring decisions based on perceptual details or matches to schemas/templates (see Johnson et al., 2003). The findings of Knott and Dewhurst also mirrored those of the present study in that the increase in false remember responses was observed with both DRM and categorized lists, indicating a further parallel between the two procedures.

In another recent study, Meade, Watson, Balota, and Roediger (2007) suggested that associative memory illusions require participants to be in retrieval mode (Tulving, 1983). They found that priming from DRM lists (as measured by a lexical decision task) is short-lived and concluded that longer-lasting effects of the sort that give rise to false recognition require participants to be given explicit retrieval instructions and recognition probes that are strongly related to the critical lures. However, while retrieval mode may be important in order for false memories to occur, the findings from the present study indicate that false memories are influenced by associations activated at study but not by associations activated at test, and that this is the case for both DRM and categorized lists.

References

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory & Language, 48,* 1–15

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.

Coane, J. H., & McBride, D. M. (2006). The role of test structure in creating false memories. *Memory & Cognition, 34,* 1026–1036.

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

Dewhurst, S.A. (2001). Category repetition and false recognition: Effects of instance frequency and category size. *Journal of Memory & Language, 44*, 153-167.

Dewhurst, S.A., & Anderson, S.J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition, 27*, 665-673.

Dewhurst, S.A., Barry, C., & Holmes, S.J. (2005). Exploring the false recognition of category exemplars: Effects of divided attention and explicit generation. *European Journal of Cognitive Psychology, 17*, 803-819.

Dewhurst, S.A., Barry, C., Swannell, E.R., Holmes, S.J., & Bathurst, G.L. (2007). The effects of divided attention on false memory depend on how memory is tested. *Memory & Cognition, 35*, 660-667.

Dewhurst, S.A., & Farrand, P. (2004). Investigating the phenomenological characteristics of false recognition for categorised words. *European Journal of Cognitive Psychology, 16*, 403-416.

Dodd, M. D., Sheard, E. D., & MacLeod, C. M. (2006). Re-exposure to studied items at test does not influence false recognition. *Memory*, *14*, 115-126.

Dunn, J.C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111*, 524-542.

Düzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.-J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Science, 94*, 5973-5978.

Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 120-128.

Gallo, D.A. (2006). *Associative illusions of memory*. New York: Psychology Press.

Gardiner, J.M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*, 309-313.

Gardiner, J.M., Konstantinou, I., Karayianni, I., & Gregg, V.H. (2005). Memory awareness following speeded compared with unspeeded picture recognition. *Experimental Psychology, 52,* 1-10.

Gardiner, J.M., Ramponi, C., & Richardson-Klavehn, A. (1999). Response deadline and subjective awareness in recognition memory. *Consciousness and Cognition, 8,* 484-496.

Gardiner, J.M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory, 10,* 83-98.

Gardiner, J.M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F.I.M. Craik (Eds), *The Oxford handbook of memory* (pp. 229-244). New York: Oxford University Press.

Howe, M.L., Wimmer, M.C., & Blease, K. (in press). The role of associative strength in children's false memory illusions. *Memory*.

Hunt, R. R., & Einstein, G. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior, 20,* 497–514.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, *114,* 3-28.

Kimball, D.R., & Bjork, R.A. (2002). Influences of intentional and unintentional forgetting on false memories. *Journal of Experimental Psychology: General, 131*, 116-160.

Kimball, D.R., Bjork, E.L., Bjork, R.A., & Smith, T. (2008). Part-list cuing and the dynamics of free recall. *Psychonomic Bulletin & Review, 15,* 296-301.

Kirkpatrick, E.A. (1894). An experimental study of memory. *Psychological Review, 1*, 602-609.

Knott, L.M., & Dewhurst, S.A. (2007). The effects of divided attention at study and test on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition, 35,* 1954-1965.

McCabe, D.P., Presmanes, A.G., Robertson, C.L., & Smith, A.D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review, 11*, 1074-1079.

McDermott, K.B. (1996). The persistence of false memories in list recall. *Journal of Memory & Language, 35*, 212-230.

Marsh, E.J., & Dolan, P.O. (2007). Test-induced priming of false memories. *Psychonomic Bulletin & Review, 14*, 479-483.

Marsh, E.J., Dolan, P.O., Balota, D.A., & Roediger, H.L., III. (2004). Part-set cuing effects in younger and older adults. *Psychology & Aging, 19,* 134-144.

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

Meade, M.L., Watson, J.M., Balota, D.A., & Roediger, H.L., III. (2007). Spreading activation and false memories in the DRM paradigm: Retrieval mode is necessary. *Journal of Memory & Language, 56*, 305-320.

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

Nickerson, R.S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in memory research. *Memory & Cognition*, *12*, 531-552.

Park, L., Shobe, K.K., & Kihlstrom, J.F. (2005). Associative and categorical relations in the associative memory illusion. *Psychological Science*, *16*, 792-797.

Reysen, M.B., & Nairne, J.S. (2002). Part-set cuing of false memories. *Psychonomic Bulletin & Review, 9*, 389-393.

Roediger, H.L., III. (1978). Recall as a self-limiting process. *Memory & Cognition*, *6*, 54-63.

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

Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. *Psychonomic Bulletin and Review*, *8*, 385-407.

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: Subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review*, *12*, 865-873.

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.

Seamon, J.G., Luo, C.R., Schlegel, S.E., Greene, S.E., & Goldenberg, A.B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory and Language*, *42*, 120-146.

Slamecka, N.J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76*, 504-513.

Smith, S.M., Gerkens, D.R., Pierce, B.H., & Choi, H. (2002). The roles of associative responses at study and semantically guided recollection at test in false memory: the Kirkpatrick and Deese hypotheses. *Journal of Memory & Language, 47*, 436-447.

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

Toglia, M.P., Neuschatz, J.S., & Goodwin, K.A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7*, 233-256.

Tse, C-S., & Neely, J.H. (2007). Semantic and repetition priming effects for Deese/Roediger-McDermott (DRM) critical items and associates produced by DRM and unrelated study lists. *Memory & Cognition, 35*, 1047-1066.

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *26*, 1-12.

Tussing, A.A., & Greene, R.L. (1997). False recognition of associates: How robust is the effect? *Psychonomic Bulletin & Review, 4*, 572-576.

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122-129.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: an updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory & Language, 50*, 289–335.

Table 1. Mean proportions of correct and false remember and know responses (with standard errors) as a function of list type and presentation order in Experiment 1.

Correct recognition				
	DRM		Categorized	
	Blocked	Random	Blocked	Random
Remember	.49 (.03)	.32 (.03)	.51 (.03)	.43 (.03)
Know	.15 (.02)	.14 (.01)	.19 (.02)	.17 (.01)
Guess	.08 (.02)	.10 (.02)	.07 (.01)	.09 (.02)
Total	.71 (.03)	.56 (.04)	.77 (.03)	.69 (.03)
Critical lures				
	DRM		Categorized	
	Blocked	Random	Blocked	Random
Remember	.33 (.03)	.15 (.02)	.20 (.03)	.08 (.02)
Know	.19 (.02)	.20 (.02)	.16 (.02)	.11 (.01)
Guess	.12 (.02)	.14 (.02)	.10 (.03)	.11 (.03)
Total	.63 (.04)	.49 (.05)	.46 (.04)	.30 (.04)
Unrelated lures				
	DRM		Categorized	
	Blocked	Random	Blocked	Random
Remember	.08 (.02)	.08 (.02)	.05 (.02)	.03 (.02)
Know	.06 (.01)	.07 (.01)	.04 (.01)	.06 (.01)
Guess	.08 (.01)	.08 (.02)	.07 (.02)	.10 (.04)

Total	.22 (.03)	.22 (.03)	.16 (.03)	.19 (.04)

Table 2. Mean proportions of correct and false remember and know responses (with standard errors) as a function of list type and encoding instructions in Experiment 2.

Correct recognition				
	DRM		Categorized	
	Read	Generate	Read	Generate
Remember	.49 (.01)	.54 (.02)	.46 (.03)	.51 (.02)
Know	.13 (.01)	.08 (.01)	.13 (.01)	.11 (.01)
Guess	.09 (.01)	.07 (.01)	.08 (.01)	.06 (.01)
Total	.71 (.02)	.68 (.02)	.67 (.03)	.68 (.02)
False recognition				
	DRM		Categorized	
	Read	Generate	Read	Generate
Remember	.29 (.03)	.42 (.03)	.11 (.01)	.17 (.02)
Know	.19 (.02)	.17 (.01)	.13 (.02)	.15 (.02)
Guess	.17 (.02)	.10 (.02)	.14 (.02)	.16 (.03)
Total	.64 (.03)	.69 (.03)	.38 (.03)	.48 (.04)
Unrelated lures				
	DRM		Categorized	
	Read	Generate	Read	Generate
Remember	.06 (.01)	.04 (.01)	.02 (.01)	.05 (.01)
Know	.08 (.01)	.07 (.01)	.04 (.01)	.06 (.01)
Guess	.15 (.03)	.11 (.02)	.05 (.02)	.04 (.01)

Total	.29 (.04)	.22 (.04)	.10 (.03)	.14 (.02)

Table 3. Mean proportions of correct and false remember and know responses as a function of list type and number of test primes in Experiment 3.

Correct recognition

Self-paced	DRM	Categorized
Remember	.41 (.04)	.49 (.04)
Know	.15 (.02)	.11 (.02)
Guess	.09 (.01)	.07 (.01)
Total	.66 (.02)	.68 (.03)
Speeded	DRM	Categorized
Remember	.28 (.02)	.31 (.02)
Know	.09 (.01)	.12 (.01)
Guess	.07 (.01)	.07 (.01)
Total	.44 (.03)	.49 (.03)

False recognition (studied lists)

Self-paced	DRM			Categorized		
	0	3	6	0	3	6
Remember	.33 (.05)	.40 (.05	.26 (.05)	.19 (.03)	.23 (.04)	.25 (.04)
Know	.23 (.04)	.14 (.03)	.21 (.04)	.15 (.04)	.11 (.02)	.05 (.02)
Guess	.10 (.02)	.11 (.03)	.11 (.03)	.12 (.02)	.11 (.02)	.10 (.03)
Total	.66 (.00)	.66 (.00)	.58 (.00)	.46 (.04)	.44 (.05)	.40 (.05)
Speeded	DRM			Categorized		

	0	3	6	0	3	6
Remember	.21 (.04)	.22 (.04)	.12 (.03)	.11 (.03)	.14 (.03)	.08 (.03)
Know	.11 (.03)	.15 (.02)	.10 (.03)	.07 (.03)	.14 (.03)	.13 (.03)
Guess	.11 (.02)	.07 (.02)	.10 (.02)	.11 (.03)	.13 (.03)	.10 (.02)
Total	.42 (.05)	.44 (.05)	.32 (.05)	.29 (.04)	.41 (.05)	.32 (.04)
False recognition (unstudied lists)						
Self-paced	DRM			Categorized		
	0	3	6	0	3	6
Remember	.05 (.02)	.10 (.03)	.07 (.02)	.03 (.01)	.05 (.02)	.05 (.02)
Know	.09 (.03)	.15 (.03)	.07 (.02)	.06 (.02)	.05 (.02)	.03 (.01)
Guess	.07 (.02)	.07 (.02)	.04 (.02)	.04 (.02)	.10 (.03)	.03 (.01)
Total	.21 (.00)	.31 (.00)	.19 (.00)	.13 (.04)	.15 (.04)	.13 (.03)
Speeded	DRM			Categorized		
	0	3	6	0	3	6
Remember	.03 (.02)	.03 (.02)	.04 (.03)	.01 (.01)	.02 (.02)	.00 (.00)
Know	.10 (.03)	.08 (.02)	.03 (.02)	.01 (.01)	.03 (.02)	.01 (.01)
Guess	.10 (.02)	.10 (.03)	.09 (.02)	.05 (.03)	.05 (.02)	.06 (.02)
Total	.23 (.05)	.20 (.04)	.16 (.04)	.07 (.04)	.10 (.03)	.06 (.02)

Unrelated lures		
Self-paced	DRM	Categorized
Remember	.05 (.01)	.04 (.01)

Know	.07 (.01)	.04 (.01)
Guess	.05 (.01)	.04 (.01)
Total	.17 (.02)	.11 (.03)
Speeded	DRM	Categorized
Remember	.02 (.01)	.01 (.01)
Know	.04 (.01)	.02 (.01)
Guess	.06 (.01)	.06 (.02)
Total	.13 (.02)	.09 (.03)

Table 4. Mean correct and false recall rates (with standard errors) as a function of list type and test cues.

	Uncued	Cued
Correct recall		
DRM	.49 (.02)	.41 (.02)
Categorized	.61 (.03)	.51 (.03)
False recall		
DRM	.42 (.03)	.33 (.05)
Categorized	.13 (.03)	.15 (.03)