The effects of mediated word lists on false recall and recognition

Mark J. Huff · Keith A. Hutchison

Abstract False memory effects were explored using unrelated list items (e.g., *slope*, *reindeer*, *corn*) that were related to mediators (e.g., *ski*, *sleigh*, *flake*) that all converged upon a single nonpresented critical item (CI; e.g., *snow*). In Experiment 1, participants completed either an initial recall test or arithmetic problems after study, followed by a final recognition test. Participants did not falsely recall CIs on the initial test; however, false alarms to CIs did occur in recognition, but only following an initial recall test. In Experiment 2, participants were instructed to guess the CI, followed by a recognition test. The results replicated Experiment 1, with an increase in CI false alarms. Experiment 3 controlled for item effects by replacing unrelated recognition items from Experiment 1 with both CIs and list items from nonpresented lists. Once again, CI false alarms were found when controlling for lexical characteristics, demonstrating that mediated false memory is not due simply to item differences.

Keywords Mediated priming · False memory · Testing effect · Recognition

In the Deese/Roediger–McDermott (DRM) paradigm, participants are presented with word lists of associatively related items (e.g., *bed*, *rest*, *tired*, *dream*, *slumber*) that all converge upon a single, nonpresented critical item (e.g., *sleep*; Deese, 1959; Roediger & McDermott, 1995).

M. J. Huff (✉)
Department of Psychology, University of Calgary,
Calgary, Alberta, Canada
e-mail: mjhuff@ucalgary.ca

K. A. Hutchison
Department of Psychology, Montana State University,
Bozeman, Montana

Participants demonstrate robust false memories for this critical item (CI) as having been presented during a previous study episode, even though the CI itself was not studied. This effect has been demonstrated across several types of testing conditions, including recall, recognition (Gallo, Roediger & McDermott 2001; Roediger & McDermott, 1995), implicit stem completion (McDermott, 1997), and part-list cuing (Bäulm & Kuhbandner, 2003), as well as across different presentation manipulations, including duration (see Gallo, 2006, 2010, for reviews), number of study items (Robinson & Roediger, 1997), and nonconscious processing (Cotel, Gallo & Seamon 2008). Taken together, these studies demonstrate that the DRM false memory effect may only be slightly moderated by differences in testing and presentation procedures; however, the cause of these false memories may be contingent on the associative characteristics of the list items themselves.

Evidence for the importance of association between the studied list items and the CI is not new. Deese (1959) demonstrated that backward associative strength (BAS) from the list items to the CI was strongly correlated with the probability that a participant would intrude the critical item during recall. Roediger, Watson, McDermott and Gallo (2001) conducted a regression analysis to determine the list characteristics that best predict false memory of the CI. The results indicated that across variables, BAS was the best predictor of critical intrusions on DRM lists, demonstrating the importance of direct semantic association for false memories (cf. Hutchison & Balota, 2005). Although the DRM effect highlights the importance of associative strength in the production of memory errors, only direct relations between the CI and the list items have been examined to date. The purpose of the present study was to examine false memory effects by using study items that are indirectly associated with a nonpresented CI via nonpresented mediators.

Evidence for the influence of indirectly associated words on target recognition has been demonstrated through research examining semantic priming. The *semantic priming effect* refers to the finding that participants are faster to respond to a word (e.g., *table*) if it was preceded by a related word (e.g., *chair*) than if it was preceded by an unrelated word (e.g., *pencil*; Meyer & Schvaneveldt, 1971; Neely, 1977). In mediated priming experiments, participants respond to word items (e.g., *beach*) that are indirectly related to a preceding prime (e.g., *box*) through a nonpresented mediator (e.g., *sand*). The main finding of these experiments is that, similar to semantic priming, participants are faster at responding to targets following primes related via a mediator (e.g., *beach–box*) than to targets following unrelated primes (e.g., *fence–box*). Mediated activation is typically explained through the spreading activation model (Collins & Loftus, 1975). According to spreading activation, when a concept becomes activated, this activation not only spreads to directly related concepts, but also to indirectly related concepts (Balota & Lorch, 1986; Chwilla & Kolk, 2002; McNamara & Altarriba, 1988). Priming effects occurring over one or more associative steps have, until recently, been considered some of the best evidence for a spreading activation theory of semantic priming (Hutchison, 2003); however, more recent data have suggested that the spreading activation account alone may not be able to explain all mediated effects, with some results supporting an additional integration or semantic matching account (cf. Hutchison & Davis, 2010; Jones, 2010).

A critical question when researching episodic false memory effects is whether the increase in false memory for nonpresented items occurs because of lexical associative activation from the studied items to the CI or because of *gist extraction*—the extraction of meaning from the list items that facilitates processing of the nonpresented CI. The notion that the organization of conceptual representations plays a large role in memory performance is well established. For example, Bousfield (1953) reported that when participants studied a word list of random items, subsequent free recall tests were often organized such that items would be grouped by a related category. Similarly, Miller (1956) suggested that combining list items in categorizable "chunks" could increase the relatively limited capacity of short-term memory. Categorically related stimuli also tend to influence recall performance, with participants recalling a greater number of items when studying related word lists as opposed to unrelated word lists (cf. Rabinowitz, Craik & Ackerman, 1982). These effects occur even when item characteristics such as frequency, word length, concreteness, familiarity, and imageability are controlled (Huff, Meade & Hutchison, under review). Taken together, semantic organization appears to

improve memory; however, this organization could come at a cost if individuals produce a plausible, semantically similar item that was not presented during study.

Fuzzy-trace theory (Brainerd & Reyna, 1990) accounts for such organization-based false memories by reference to *verbatim* and *gist* representations. For example, when an individual is presented with the word "tired," a verbatim representation is stored, consisting of specific presentation details of the word itself, as well as a gist representation containing the overall meaning of the word, "lack of energy." Because false memories would not possess a verbatim representation, the manifestation of false memories is presumably due to the stored gist representation being retrieved at recall. The powerful DRM effect emerges because the meaning of the nonpresented CI is similar to several gist representations stored during encoding of the list (Brainerd & Reyna, 2002).

A strength of the fuzzy-trace theory is its ability to predict long-term false memories. Specifically, DRM false memories can be long-lasting, with recall and recognition effects persisting months after study (Seamon et al., 2002). Over a delay, verbatim traces fade at a faster rate than the gist traces. These persisting gist traces are responsible for the production of the CI. In contrast, implicit associative activation tends to be short-lived following a single item (Neely, 1977) or even an entire DRM list (Meade, Watson, Balota & Roediger 2007; but see Tse & Neely, 2005), predicting an absence of DRM false memories following a delay. In order to account for this discrepancy, Meade et al. (2007; see also Hutchison & Balota, 2005) hypothesized that participants engage in retrieval-mode processing (Tulving, 1983) after a delay, during which the deliberate attempt to retrieve items from episodic memory causes reactivation of the associative network encoded during study.

The greatest problem in trying to separate the contributions of gist extraction versus associative spreading activation when studying DRM false memories is that BAS and similarity in meaning are highly confounded. Associated items share a variety of semantic relations, and strong associates tend to have a larger overlap in semantic features (Hutchison, 2003). As a result, higher-BAS lists may be more thematically consistent with the CI. In addition, Brainerd, Yang, Reyna, Howe and Mills (2008) found that high-BAS lists tend to have CIs that have been independently rated as more familiar and meaningful in Toglia and Battig's (1978) semantic word norms. Thus, it may have been these semantic characteristics of the critical items themselves, rather than BAS, that led to higher false memories in such lists.

In an attempt to separate these factors, Hutchison and Balota (2005) held BAS constant, while manipulating the number of meanings present in a study list. One set of items used were lists similar to those used in typical DRM

experiments, in which all items converge upon one non-presented CI that has a single meaning. The other set of items were lists that each converged upon a single homograph CI (e.g., *fall*), but one half of the items converged upon one meaning of the homograph ("stumble"), and the other half converged upon the other meaning ("autumn"). The authors hypothesized that if false memories were due to gist extraction, false memories would increase along with increased list length for DRM-type lists, but not for homograph lists in which the additional items related to the second meaning. This is because adding the second meaning on homograph lists would conflict with the gist representation already formed from the first meaning. Conversely, if false memories were due to associative activation, the addition of items on DRM-type lists and the second meaning on homograph lists would increase false memories equally, because the corresponding associative strength would increase equally for both list types. The results indicated that when a second homograph meaning was studied, the increase in CI intrusions was identical to the increase in CIs when only one meaning was studied (i.e., DRM lists), supporting the associative activation over the gist extraction account.

Hutchison and Balota (2005) also manipulated the consistency of meaning between list items by using homograph word lists that either were blocked or alternated by meaning. Thematic consistency and gist identifiability should have been greater when meanings were blocked, relative to when meanings were alternated. However, cross-experimental comparisons revealed that participants falsely recalled CIs at similar rates, regardless of whether homograph lists were blocked or alternated by meaning, despite a drop in veridical recall for alternated lists. Further, when participants were asked to judge CIs on how similar in meaning they were to items in a studied list, DRM items were judged as more similar than homograph items, despite no differences in false memories between the two item types. Thus, the authors concluded that the thematic consistency of meaning in studied lists does not appear to play a role in the occurrence of DRM false memories.

The present set of experiments also sought to separate gist- and association-based false memory representations, by using lists in which no relations were shared between the items, nor did the items share a common gist representation; instead, items were only indirectly related, through mediators that converged upon a single, nonpresented CI. If false recognition effects can occur in the absence of thematic consistency between items, gist extraction processes could not then account for false memory effects. In essence, by presenting participants with word items that are only related to a nonpresented CI through mediators, these experiments effectively tested associative-based false memory processes in the absence of gist-based strategic relational processing.

In the experiments reported here, we presented participants with study lists composed of word items directly associated with mediators that, in turn, all converged upon a single nonpresented CI. There was little to no association between list items themselves, and no direct association between a study item and the CI. In Experiment 1, participants studied items at two durations, followed by an immediate recall test or by arithmetic problems. At the end, participants completed a final recognition test over all studied lists. The critical research questions for Experiment 1 were whether there would be an increase in false alarms for CIs relative to new unrelated items and whether these false alarms would differ as a function of presentation duration and initial recall test. In Experiment 2, we examined whether participants might become consciously aware of the mediated CIs during study, but choose not to report them during recall. In Experiment 3, we explored the possibility that increased false alarm rates for mediated CIs might be due solely to lexical characteristics within the items themselves. Results across all experiments demonstrated strong support for mediated false memory effects, and these effects were not due to conscious processing of CIs or item differences that occurred on the final recognition test. Furthermore, these experiments demonstrated support for association-based processes underlying false memory effects.

Experiment 1

Method

Participants Forty male and female undergraduates at Montana State University participated in exchange for partial completion of a research requirement for an introductory psychology course. All were native English speakers with normal or corrected-to-normal vision.

Materials Mediated triads (e.g., *summer–winter–snow*) were used to create 24 word lists (presented in the Appx.) composed of 8 items each using the Nelson, McEvoy and Schreiber (1999) word association norms. The mediated items were taken from Balota and Lorch (1986). These lists were constructed by generating 8 items (e.g., *ski, sleigh, flake, . . .*) with high associative strength to a target word (e.g., *snow*). Using this list of 8 mediators, study list items were created by taking the word with the greatest associative strength to the mediator that did not also have an association to the target (e.g., *slope, reindeer, corn, . . .*), producing unrelated 8-item word lists that were directly associated with a list of mediators that then converged upon a single nonpresented target. During study, these lists were blocked together in pairs to create 12 word lists, each 16

items in length. Therefore, for each studied list, participants could potentially recall or recognize two critical non-presented targets.

The final recognition tests were composed of 1 critical item, 2 list items, and 3 unrelated nonpresented words from each 8-item study list, resulting in a total of 144 recognition items. These items were split into two groups to form two 72-item recognition tests that were each presented at the end of the first and second halves of the experiment. Unrelated test items were selected using the English Lexicon Project database (Balota et al. 2007) and were matched to list items based on average word length and logarithmic frequency of occurrence in the English language via the Hyper Analogue to Language database (Lund & Burgess, 1996).

Procedure All stimuli were presented and the recognition test was completed using E-Prime software (Schneider, Eschman & Zuccolotto 2002). Each participant was tested individually and was seated approximately 60 cm away from a VGA monitor. Participants were presented with instructions on the monitor that were read by the experimenter to the participant. Participants were informed that they would be presented with a list of words at either a 500-ms or a 3,000-ms duration and that their task would be to attend to each item in preparation for a later memory test. The interstimulus interval (ISI) for each list was 500 ms. The 500- and 3,000-ms list durations were counterbalanced across participants. Following the presentation of each word list, participants then completed either a recall test on a separate sheet of paper or arithmetic problems for 60 s. Following six study–test/arithmetic trials, participants completed a 72-item old/new judgment recognition test in which they were informed to respond quickly but without compromising accuracy.

Following the recognition test, participants repeated the same procedure with six new study lists, followed by another final recognition test. Following the completion of the second recognition test, all participants were fully debriefed and awarded credit for participation. A typical experimental session lasted approximately 40 min.

Design The experiment utlized a 2 (presentation duration: 500 vs. 3,000 ms) × 2 (initial test: recall vs. arithmetic) × 3 (recognition item type: studied vs. critical vs. unrelated) mixed design with duration and item type as within-subjects variables and initial test as a between-subjects variable.

Results and discussion

For all results reported, statistical significance was set at $p < .05$ unless otherwise noted.

Recall The proportions of correctly recalled list items, along with the proportions of critical intrusions recalled, for the 3,000- and 500-ms presentation durations for the 20 participants who completed recall tests are presented in Table 1. Analysis of the proportions of correctly recalled items demonstrated that participants recalled $20 \pm 8\%$ ($\pm$ = 95% confidence interval) more words from lists presented for a 3,000-ms duration than from lists presented for a 500-ms duration, $t(19) = 9.90$, $SEM = .02$. However, participants did not recall the mediated critical items at either presentation duration ($t < 1$).

Recognition The proportions of studied list items, critical items, and unrelated items given an "old" response are presented in Table 2. Two separate 2 (recall: prior recall vs. no prior recall) × 2 (duration: 3,000 vs. 500 ms) mixed factorial ANOVAs were used to examine duration and initial recall differences on hit rates and false alarms. For hit rates on studied items, the main effects of both duration and initial test were not significant (both $Fs < 1$); however, a significant Duration × Initial Test interaction was found, $F(1, 38) = 4.14$, $MSE = .01$. Post-hoc analyses demonstrated that initial testing marginally increased hit rates by $7 \pm 9\%$ for list items presented at 3,000 ms, $t(38) = 1.74$, $SEM = .03$, but had no effect ($0 \pm 11\%$) when items were presented for 500 ms.

A 2 (recall) × 2 (duration) × 2 (item type: critical vs. unrelated) ANOVA was used to test for false recognition for mediated CIs. As was found for hit rates, there were no main effects for either initial testing or duration (both $Fs < 1$); however, a significant main effect of item type was found, $F(1, 38) = 11.57$, $MSE = .01$, demonstrating that false alarms to critical nonpresented items were $11 \pm 9\%$ greater than false alarms to new unrelated items. Additionally, this main effect was qualified by a significant Initial Testing × Item Type interaction, $F(1, 38) = 5.70$, $MSE = .01$. Pairwise comparisons revealed that participants were $10 \pm 5\%$ more likely to produce false alarms to critical than to unrelated items after an initial recall test, $t(19) = 3.50$, $SEM = .03$, but showed no such elevation in false alarms to critical items (a $2 \pm 5\%$ difference) following an arithmetic task ($t < 1$). The three-way interaction between duration, recall, and item type failed to reach significance, $F(1, 38) = 1.06$, $p = .31$.

Interestingly, mediated false memory effects on recognition tests only occurred if participants completed an

Table 1 Recall data of mediated items in Experiment 1 as a function of presentation duration (3,000 or 500 ms)

	3,000 ms	500 ms
Veridical recall	.47 (.09)	.27 (.05)
Critical intrusions	.00 (.03)	.00 (.03)

Standard deviations are listed in parentheses.

Table 2 Recognition hits (studied items), recognition false alarms (critical nonpresented items and new unrelated items), and corrected critical recognition (critical items minus nonpresented items) of mediated items in Experiment 1 as a function of a prior recall test and presentation duration (3,000 or 500 ms)

	3,000 ms	500 ms
Studied Items		
Prior recall	.75 (.16)	.70 (.19)
No prior recall	.67 (.12)	.70 (.14)
Critical Items		
Prior recall	.31 (.27)	.29 (.27)
No prior recall	.23 (.13)	.26 (.12)
Unrelated Items		
Prior recall	.19 (.23)	.20 (.23)
No prior recall	.22 (.09)	.22 (.10)
Corrected Critical Recognition		
Prior recall	.12 (.19)	.09 (.16)
No prior recall	.01 (.12)	.04 (.10)

Standard deviations are listed in parentheses.

immediate recall test after the presentation of the study list. This suggests that participants completing initial recall may have engaged in a type of association-based "retrieval mode" in which they reactivated associative pathways that had been created during encoding (Meade et al., 2007; Tulving, 1983). Such retrieval-mode processing could have then led to implicit or explicit activation of the mediated CI. For instance, as an individual encodes associatively related information such as a DRM list, the individual creates a network of associatively related concepts that are encoded in addition to the individual items themselves (Anderson & Bower, 1972). As individuals complete a memory search, they retrieve not only the specific studied items, but also the network of associates that were active during encoding (Anderson, 1983; Hutchison & Balota, 2005; Meade et al., 2007). According to Anderson's (1983) ACT model, when strong associates of recalled items are activated during retrieval, preexisting associative connections between these items and studied items can give the illusion that such items were studied alongside list items. When sufficient connections exist, an individual may have enough memory evidence to report that an item was presented during an earlier study session.

The use of such a retrieval mode may be particularly relevant to the present experiment, because as participants begin to recall a set of mediated items, they may in turn activate the nonpresented mediators that all converge upon the nonpresented CI, resulting in greater levels of false memory. Therefore, it is possible that the original associative network established during encoding might be strengthened (and perhaps expanded upon) when tested. This may, in turn, produce memory errors that are

associatively related. In fact, Roediger and McDermott (1995) demonstrated that participants who completed an initial recall test had an increase in false alarms to the associatively related CI on a final recognition test relative to participants who completed arithmetic problems. Although this increase in CI false alarms was strongly related to falsely recalling the CI during initial testing, it is possible that initial testing also increased the associative relationships between the items, thereby inflating false alarms on the subsequent recognition test. This process may be particularly relevant for retrieval mode, in which the associations become reactivated; the completion of an initial test may increase the associative network between study items, resulting in a greater false alarm rate for indirectly associated lures through a nonpresented mediator.

In summary, the present experiment provides evidence for mediated false memories. Presenting participants with unrelated study lists all sharing nonpresented mediators that converge upon a single CI appears to inflate false alarm rates for the CI relative to an unrelated nonpresented item. Furthermore, these data support the notion that associations that occur between list items and nonpresented mediators give rise to false memory effects. Since the list items are not thematically consistent, only associative processes could explain these effects. Interestingly, this false memory effect appears to occur only on a recognition test, since the CI did not intrude for participants taking a recall test; however, it is possible that participants consciously generated the mediated CI during initial recall, but chose not to report it because they knew it was not presented during study. This would suggest that participants later had a source failure during recognition by falsely reporting the CI that had been consciously identified, yet rejected, during initial recall. This possibility was examined in Experiment 2.

Experiment 2

The primary purpose of Experiment 2 was to examine the possibility that participants might consciously generate the nonpresented CIs during study, but choose not to report them during initial recall. Additionally, Experiment 2 sought to replicate the mediated false recognition effect obtained in Experiment 1.

Method

Stimuli and procedure The same procedure and stimulus word lists used in Experiment 1 were used for Experiment 2, with the following exceptions. First, due to the lack of duration differences found in Experiment 1, the presentation duration for all list items was 2,000 ms (500-ms ISI). Second, participants were instructed to guess the critical

nonpresented item following presentation of each word list. Specifically, participants were informed that the first and the last eight items of each list converged upon a single nonpresented word, and they were to try to guess this item. Following each list presentation, participants were given 60 s to report a convergent item for each of the two presented eight-item sublists. This guessing procedure took the place of the recall test and arithmetic problems used in Experiment 1. Participants were required to provide one response for each sublist, resulting in a total of two guesses for each list studied. In addition, participants were also required to provide a confidence rating for each guess, using a four-point scale (4 = *very confident that this was the converging item*, 3 = *somewhat confident that this was the converging item*, 2 = *somewhat not confident that this was the converging item*, 1 = *very confident that this was not the converging item*). As in Experiment 1, participants then completed the identical final recognition test. A typical experimental session lasted approximately 40 min.

Participants Twenty male and female Montana State University undergraduates participated in exchange for partial completion of course credit for an introductory psychology class. All participants were native English speakers with normal or corrected-to-normal color vision.

Results and discussion

CI guessing The proportions of correctly guessed critical items were calculated by taking the number of correctly guessed critical items divided by the total number of guesses provided. While scoring guess rates, a liberal criterion to score CI identification was used, since participants would often report an item that was synonymous to the nonpresented CI but not report the exact item (i.e., reporting "dim" when the correct CI was "dark"). This criterion was adopted in order to assess the possibility that the participant was conceptually aware of the CI during list presentation. It is important to note that this liberal criterion was not used for the recall tests in our other experiments. As is standard in the DRM literature, items in the recall tests were judged as recalled only if the word itself was recalled, not if a related word was recalled.

Correct guess rates were analyzed by comparing the proportion to zero using a single-sample *t* test. As can be seen in Table 3, participants were significantly able to guess the CI about 5 ± 2% of the time, $t(19) = 5.00$, $SEM = .01$. Therefore, when participants were instructed about the CI prior to the presentation of each list, the CI (or a similar concept) could occasionally be successfully identified. However, even under such instructional conditions, correct identification was a rare occurrence.

Table 3 Proportions of critical items correctly guessed after presentation of mediated word items, as well as recognition hits (studied items), recognition false alarms (critical nonpresented items and new unrelated items), and corrected critical recognition (critical items minus nonpresented items) of mediated items in Experiment 2

Variable	
Proportion of guessing critical item	.05 (2.37*)
Recognition Test	
Studied items	.80 (.13)
Critical items	.43 (.15)
Unrelated items	.27 (.12)
Corrected critical	.15 (.12)

Standard deviations are listed in parentheses. Presentation duration was 2,000 ms and N=20.

*Mean confidence rating, 1–4 scale.

Confidence ratings were also examined for CIs that were correctly and incorrectly guessed; the mean rating for correctly guessed CIs is also reported in Table 3. Participants reported slightly greater confidence for correctly guessed CIs ($M = 2.37$) than for incorrect items ($M = 1.83$); however, this difference was only marginally significant, $t(31) = 1.78$, $SEM = 0.31$, $p = .09$. In terms of the scale used to make these judgments, participants were not very confident that their reported guess was correct, and this was similar for both correct and incorrect guesses. Taken together, participants were very seldom able to guess the nonpresented CI, and even these rare correct guesses were made with little confidence. This pattern suggests that the presentation of these stimuli did not bring the mediated CI to mind consciously during initial study.

Recognition Proportions of studied list items, critical items, and unrelated items given an "old" response are presented in Table 3. In order to examine recognition differences, a repeated measures ANOVA was used to assess item type differences. A main effect of item type was found, $F(2, 38) = 120.67$, $MSE = .01$, demonstrating that differences in the proportions of "old" responses depended on the types of test items. Pairwise comparisons revealed that correct responses to list items were 53 ± 8% greater than false alarms to unrelated items, $t(19) = 13.56$, $SEM = .03$, and 37 ± 8% greater than false alarms to CIs, $t(19) = 9.85$, $SEM = .03$. Critically, differences were also found between false alarms for CIs and unrelated items, such that participants were 15 ± 6% more likely to produce a false alarm to CIs than to unrelated items, $t(19) = 5.78$, $SEM = .03$.

One interesting finding from the recognition data in the above experiment is the replication of the initial-testing condition from Experiment 1 using different instructions. According to the data, it appears that instructing participants to attempt to guess the nonpresented CI after the presentation

of each mediated list produced a significant false recognition effect similar to when participants completed an initial recall test (cf. Table 2). This pattern suggests that the processes involved in testing and attempting to guess the nonpresented item may be similar, in that they both increase later false recognition. Both completion of a recall test and guessing may have caused participants to engage in a retrieval mode in which the associative activation that was established during list presentation became reactivated during test.

Experiment 3

One potential issue with the above experiments is that the high levels of false alarms to critical items, relative to unrelated items on the recognition test, may have actually been due to differences in item characteristics. Although unrelated items were carefully selected to control for differences in frequency and word length relative to list items, these unrelated items were not matched to the critical items. Characteristics of certain items have been shown to influence memory performance in both recall and recognition (Cortese, Khanna & Hacker 2010), which may mask any performance differences. Indeed, Neely, Johnson, Neill and Hutchison (1999) demonstrated recognition false alarm differences between critical items and list items even following unrelated study lists, suggesting that part of the traditional "false memory" effect is due simply to differences in item characteristics between the list items and critical items. To control for such potential item confounds, the recognition test was restructured so that critical items from lists that were not viewed by participants were used as unrelated test items during recognition.

Method

Stimuli and procedure The stimuli and procedure for Experiment 3 were identical to those in Experiment 1, except that all list items were presented for 2,000 ms (500-ms ISI). First, participants only viewed half of the word lists (six 16-item word lists), immediately followed by a 60-s recall test or arithmetic problems, followed by a final recognition test. Additionally, the recognition test was modified from earlier experiments to include different test items, such that participants were tested with one critical nonpresented item from a presented list, one studied list item, one critical item from a nonpresented list, and one list item from a nonpresented list for each of the 12 presented word lists, resulting in a recognition test of 48 total items. A typical experimental session lasted approximately 20 min.

Participants Seventy male and female Montana State University undergraduate students participated in exchange for partial course credit for an introductory psychology class. All participants were native English speakers with normal or corrected-to-normal color vision.

Results and discussion

Recall The proportions of correctly recalled list items, along with the proportions of critical intrusions recalled, for the 35 participants who completed the recall test are presented at the top of Table 4. A single-sample *t* test was completed on the proportion of critical intrusions to determine whether the intrusion rate differed significantly from zero. As in Experiment 1, participants did not intrude CIs during recall ($M=.00$), and this proportion did not differ significantly from zero, $t(34) = 1.445$, $SEM = .01$, $p = .16$.

Recognition The proportions of studied list items, critical items from studied lists, list items from nonpresented lists, and critical items from nonpresented lists are presented in Table 4. First, to examine differences between the two types of list items, a 2 (recall: prior recall vs. no prior recall) × 2 (item type: presented list items vs. nonpresented list items) mixed ANOVA was used to examine the proportions of "old" responses given to test items. A main effect of item type was found, $F(1, 68) = 562.37$, $MSE = .02$, demonstrating that participants were $57 \pm 5\%$ more likely to correctly respond "old" to presented list items than to nonpresented list items. The main effect of recall failed to reach significance, $F < 1$,

Table 4 Recall proportions for studied items and nonpresented critical items, as well as recognition hits (studied items), recognition false alarms (critical nonpresented items, list items from nonpresented lists, and critical items from nonpresented lists), and corrected critical recognition (critical items minus nonpresented critical items and nonpresented list items) of mediated items in Experiment 3 as a function of prior guessing recall

	Prior Recall (N=35)	No Prior Recall (N=35)
Recall Test		
Correct recall	.35 (.10)	
Critical intrusions	.00 (.03)	
Recognition Test		
Studied items	.77 (.17)	.73 (.19)
Critical items	.27 (.21)	.30 (.18)
Nonpresented list items	.16 (.14)	.20 (.13)
Nonpresented list critical items	.20 (.18)	.25 (.17)
Corrected Recognition		
List items	.60 (.20)	.53 (.19)
Critical items	.07 (.19)	.05 (.19)

Standard deviations are listed in parentheses. Presentation duration was 2,000 ms.

as did the interaction between item type and recall, $F(1, 68) = 2.69$, $MSE = .02$, $p = .11$, demonstrating that differences in list items were not contingent upon whether or not participants completed an initial recall test. The lack of differences between recall conditions does not replicate the testing differences in Experiment 1, but the data are numerically in the same direction.

Critically, to determine differences between "old" responses between CIs from presented and from nonpresented lists, a 2 (recall: prior recall vs. no prior recall) × 2 (item type: presented list critical items vs. nonpresented list critical items) mixed ANOVA was used. Importantly, a main effect of item type was found, $F(1, 68) = 7.96$, $MSE = .02$, demonstrating that "old" responses given to CIs from presented word lists were 6 ± 5% greater than "old" responses given to CIs from nonpresented lists, confirming that mediated false memory effects were not simply due to item differences between CIs. Somewhat surprisingly, however, the main effect of recall failed to achieve significance ($p = .31$), since the completion of a recall test did not elevate the probability of "old" responses to CIs. This pattern of results is inconsistent with Experiment 1, although similar to the pattern for list items; the pattern is also numerically consistent with Experiment 1. The interaction between item type and recall failed to reach significance, $F < 1$.

One additional finding of interest in the present study is the difference in false alarm rates between the two types of unrelated items from nonpresented lists. As mentioned above, the impetus for this experiment was to determine whether false memory effects could potentially be masked by item differences found between critical items. Although data from the present experiment effectively demonstrate that item effects alone cannot account for increased false recognition of CIs, there was evidence for an influence of item characteristics on "old" responses to nonpresented stimuli. Specifically, the participants were 4 ± 2% more likely to incorrectly respond "old" to nonpresented critical items than to nonpresented list items, $t(69) = 2.20$, $SEM = .02$, which suggests that false alarm differences may be due in part to the lexical characteristics of the items themselves, inducing participants to erroneously report having studied the item previously.

General discussion

Several findings from the present series of experiments were of particular importance to DRM false memory research. First, as shown in all three experiments, false memory effects were demonstrated using list items that were not directly related to a nonpresented CI, but instead were related to several mediating items that all converged upon a single, nonpresented item. Second, in Experiments 1 and 3, CIs were not reported when

participants completed a recall test, and further, these CIs were not guessed even when participants were explicitly informed about the relationship between the word list and the nonpresented item in Experiment 2, suggesting that the structure of mediated word lists did not allow for conscious generation of the CI. Most importantly, false memory effects did occur, but only on a final recognition test, where participants were more likely to produce false alarms to the nonpresented CI than to unrelated items. Finally, this effect occurred even when item effects were controlled.

This study is the first to report an increase in false alarms for critical nonpresented items using word lists that contain no associations between list items themselves and no direct association between study items and the CI. We suggest that the cause of the mediated false memories with these lists is due to spreading activation processes. Specifically, as participants are studying a word list, they not only encode each specific item, but also encode the associative semantic network, which includes related concepts. By using mediated study items that are indirectly related to a single CI, this CI also becomes encoded through association. When an individual becomes actively engaged in retrieval-mode processes through which the specific studied items become reactivated along with the associative network created during encoding, this reactivation gives rise to false memories that do not share a direct association with studied items.

Two potential moderators that were demonstrated to have unreliable effects in inflating or deflating mediated false memories were presentation duration and implied warnings. In Experiment 1, no differences were found between presentations of 3,000 and 500 ms, which suggests that mediated spreading activation processes occur relatively quickly. In Experiment 2, participants were informed of the experimental purpose followed by a final recognition test. These instructions made participants aware of the relationship with the CI, which could have been an implied warning. Nonetheless, false alarms to CIs were elevated in this experiment, similar to the initial-testing condition in Experiment 1. It is possible, however, that a more explicit warning of the false memory effect and that nonpresented CIs might be presented on the recognition test could have differentially affected false memories, and participants may then have been able to effectively monitor the source of the activated CI, thus reducing false alarms.

The mediated nature of the list items and CIs provided evidence that the associative network created during encoding was reactivated during retrieval, which elevated false alarm rates to CIs relative to control items. Furthermore, this retrieval-mode process was not sensitive to the quality of the original memory for the specific study items, since false alarms to the activated CIs were the same across study durations in Experiment 1. However, encoding durations did play a role in both correct recall and hit rates on final

recognition tests, since these measures were diminished with less encoding time. Therefore, it is probable that mediated association and encoding duration influence independent processes. Although this possibility was not a focus in the present study, the dissociation between hit rates and false alarms in regard to encoding durations is interesting.

Additionally, this study found item differences in false alarms between nonpresented CIs and nonpresented unrelated items that were initially used for control items. Experiment 3 demonstrated that not only were false alarms to CIs from studied lists reliably greater than false alarms to CIs from nonstudied lists, but that false alarms to CIs from nonstudied lists were also reliably greater than false alarms to unrelated items from the nonpresented lists. This pattern demonstrates that specific characteristics of the items used can differentially influence the probability at which participants respond "old" (cf. Cortese et al., 2010, for a review of item characteristics and their tendency to influence hit rates and false alarms on recognition tests).

Furthermore, these experiments effectively examined the associative activation and gist extraction hypotheses of episodic false memory. Our word lists were engineered such that individual items did not share a common gist theme, but instead were associatively related to mediators that were then associatively related to the nonpresented CI. If gist extraction processes were to play a role in episodic false memory, then thematic consistency would have to be present. In all three experiments, we found significant false recognition for the CI relative to control items using lists lacking a common theme. Therefore, only the associative activation account could successfully predict this pattern. These data support findings reported by Hutchison and Balota (2005) and Howe, Wimmer, Gagnon and Plumpton (2009).

It is important to note that, although these data support associative activation, they cannot effectively rule out gist extraction processes when a consistent theme *is* present. For instance, when a consistent theme between study items is available, both associative processes and gist extraction could be in effect. In order to entirely rule out gist extraction processes, researchers would have to demonstrate a lack of false memories using stimuli that are thematically consistent but that lack any associative relationships, either direct or indirect. This endeavor of creating stimuli possessing a consistent theme and no associative relationships would be difficult to accomplish, if not impossible. Therefore, these experiments effectively demonstrate that false memories can occur in the absence of gist extraction through associative activation processes, but cannot entirely rule out gist extraction when themes between study items are present.

Recent studies have also attempted to partial out the contributions of gist retrieval and associative mechanisms. For instance, Brainerd et al. (2008) reexamined the variables used in Roediger et al.'s (2001) regression analysis to predict false memories in the DRM paradigm, and they added several new semantic variables for both studied items and CIs. Using a factor analysis, the semantic variables of familiarity and meaningfulness loaded highly with false recall and false alarms for CIs on the same factor, which is consistent with the notion that intrusions and false alarms of the CI involve the retrieval of semantic information. However, BAS also loaded highly on this factor, once again highlighting the importance of associative relationships between studied items and CIs for this illusory effect. Because association strength in association norms is confounded with semantic similarity (Hutchison, 2003), the authors suggested that this associative pattern may have been due to the semantic relations between study items and the CI rather than to associative connections.

However, the factor-analytic approach used by Brainerd et al. (2008) presents a problem for their strong conclusions. The rationale for the use of a factor analysis was to reduce the complications of muliticollinearity that potentially occur with regression analyses; however, this method restricted the ability to determine the individual contributions of specific factors, and no effort was made to examine the effects of BAS while controlling for the CI semantic factors of familiarity/meaningfulness, or vice versa. In contrast, a regression analysis would have been more likely to identify the relative and unique contributions of each variable in predicting false memory.

Another purpose of the present study was to examine how completing an initial recall test could influence false alarm rates on a final recognition test. If an initial test embellishes upon and strengthens the associative network created during encoding, one would expect that mediated CI false alarms would increase when participants completed the initial recall test rather than arithmetic problems. Evidence supporting this rationale was found in Experiment 1, in which inflated CI false alarms were demonstrated by participants completing an initial recall test; however, in Experiment 3, no significant difference was found between recall conditions. A 2 (recall: prior recall vs. no prior recall) × 2 (item type: critical vs. unrelated items) mixed ANOVA was used to examine the overall effect of initial testing, collapsing across Experiments 1 and 3. Nonpresented critical items were used as unrelated items for Experiment 3. A main effect of item type was found, $F(1, 108) = 16.04$, $MSE = .02$, demonstrating that false alarms to CIs were $7 \pm 3\%$ greater than false alarms to unrelated items; however, the interaction between recall and item type failed to reach significance, $F(1, 108) = 1.84$, $p = .18$, demonstrating that the completion of an initial test did not increase false alarms to CIs when collapsing across experiments. This pattern is likely due to the lack of initial-test differences in the results of Experiment 3, which was not altered when combined with the significant testing difference

found in Experiment 1. Future research will need to explore whether this effect is truly conditionalized on the completion of a prior recall test or attempting to guess the CI after study.

While the influence of a recall test may affect subsequent recognition, attempting to guess the CI may also play a role in subsequent recognition performance. To explore this possibility, cross-experimental analyses examining the differences in corrected false recognition to CIs revealed that attempting to guess the CI did not elevate these false alarms relative to completing a recall test. Although participants attempting to guess CIs in Experiment 2 numerically increased CI false alarms ($M = .15$), this increase was not reliably greater than Experiment 1 ($M = .11$) or 3 ($M = .11$) when nonpresented list items were used as a control.

To the extent that guessing or completing an initial recall test influences mediated false memories, it is possible that as participants study a mediated word list, they generate some of the associatively related mediators that directly converge upon the nonpresented CI. To explore this possibility, we examined whether the nonpresented mediators intruded during recall (Exps. 1 and 3) or whether the participants guessed the nonpresented mediators during the guessing procedure (Exp. 2). In both cases, participants rarely produced these mediators, either during recall testing or when guessing (less than 1% of total responses in all experiments), and therefore it is unlikely that conscious awareness of the mediators is the cause of this increased effect.

One caveat worth noting is the relatively small false memory effects with mediated lists, relative to the robust effects found within the DRM paradigm. In fact, although the demonstration of false recognition in the absence of study themes is evidence against fuzzy-trace theory, the reduction in false memory relative to typical DRM paradigms is consistent with this theory. Specifically, fuzzy-trace theory suggests that the greater thematic consistency of DRM lists should lead to greater false memory effects than the weak ones found using the present mediated lists. Similarly, finding mediated false memory on recognition but not recall tests may also demonstrate the importance of theme availability for CI intrusions in recall. Alternatively, the strong BAS and thematic consistency in standard DRM lists might cause the critical item to come to mind consciously, whereas the study lists in the present experiment were all seemingly unrelated, creating more implicit activation of the CI. Unfortunately, however, because type of test is confounded with retention interval in the present paradigm, it is unknown whether the lack of false mediated recall is due to the type of test itself or to participants' memory strength for the presented material relative to the CI. Future studies that include a final recall test and/or immediate recognition tests would help clarify potential differential effects of mediated lists on false recall versus recognition.

Importantly, however, the associative activation hypothesis also predicts reduced false memory for mediated lists, due to the considerable reduction in associative strength found with mediated items relative to the more direct associations in DRM lists. If we examine the Roediger et al. (2001) regression data from standard DRM lists, false recall for the 8 lists with a BAS of .01 or lower was only 9%, which is greatly reduced relative to the 34% obtained for the other 46 lists. Although 9% is certainly above our present mediated rate of 0%, BAS was still above zero in those lists, and forward associative strength (from the CI to the list items) averaged .04. Similar to the recall data, the false alarm rate for the 8 weak BAS standard DRM lists was 42%, which is in between the 62% obtained for their other 46 lists and the rates in the present study (.25–.31, .43, and .27–.30 for the conditions in our Exps. 1–3, respectively).

In addition, in standard DRM lists, the CI often serves as a mediator connecting list items (e.g., bed–*sleep*–rest), allowing for more CI activation, especially when participants engage in relational processing, as they are more likely to do when expecting a recall test (Hunt & Einstein, 1981). This may occur even for weak-BAS lists (e.g., fast–*swift*–Jonathan or light–*lamp*–shade) in which the CI would not normally be the first item to come to mind on the basis of any single list item alone. This type of associative linking between list items was described in detail in Anderson and Bower's (1972, 1973) FRAN and HAM models. Within this framework, participants create associative pathways during study that connect list items to one another, and these pathways often contain related but nonpresented items. This allows associated items to become entrenched within the associative network established during study. This pathway-marking process is conceptually similar to fuzzy-trace theory's gist extraction, except that pathway marking focuses on identifying connections between two list items rather than searching for an overall *gist-across-the-list*. Importantly, either of these processes should only occur with standard, rather than mediated, DRM lists.

Finally, in the present paradigm participants were presented with one list made up of two sublists that each converged upon independent CIs. This type of presentation could have weakened either gist extraction or associative-based encoding processes. It is hoped that future studies will expand upon the present paradigm so we may understand more thoroughly the necessity of thematic consistency versus associations in producing false memory.

In sum, the present experiments are the first to examine memory errors using stimuli lacking a direct relationship to a CI. When stimuli were related to mediators that directly converged upon a nonpresented CI, the evidence demonstrated an increase in false alarms for CIs. This finding demonstrates that mediated associations can create false memory errors even in the absence of thematic gist, therefore providing evidence for associative activation processes in the formation of episodic false memories.

Appendix

Table 5 The eight-item study lists and corresponding mediators used in Experiments 1–3. CIs are listed in bold, study list are presented in normal font, and mediators are presented in italics

Stripes		**Box**		**Quiet**	
Zoo	*Zebra*	Parcel	*Package*	Still	*Silent*
Plain	*Plaid*	Temper	*Fuse*	Talk	*Whisper*
Official	*Referee*	Stamp	*Mail*	Bashful	*Shy*
Possum	*Raccoon*	Warehouse	*Storage*	Waterfall	*Peaceful*
Concrete	*Solid*	Vote	*Ballot*	Obvious	*Discreet*
Stink	*Skunk*	Pliers	*Tool*	Volume	*Loud*
Point	*Dot*	Bouillon	*Cube*	Code	*Secret*
Banner	*Flag*	Brawl	*Fight*	Patience	*Calm*
Bubbles		**Necklace**		**Smell**	
Explode	*Burst*	Ear	*Earring*	Musty	*Odor*
Kettle	*Boil*	Link	*Bracelet*	Pine	*Scent*
Chew	*Gum*	Valuable	*Jewelry*	Possum	*Skunk*
Robe	*Bath*	Giraffe	*Neck*	Cigar	*Stink*
Lather	*Soap*	Whip	*Chain*	Scratch	*Sniff*
Whistle	*Blow*	Wit	*Charm*	Cologne	*Perfume*
Tide	*Detergent*	Medal	*Gold*	Essence	*Fragrance*
Teeth	*Gums*	Prism	*Crystal*	Ban	*Deodorant*
Glass		**Drink**		**Kick**	
Frail	*Breakable*	Shrimp	*Cocktail*	Shoe	*Boot*
Pane	*Window*	Saloon	*Bar*	Goal	*Soccer*
Saucer	*Cup*	Slurp	*Sip*	Fist	*Punch*
Cookie	*Jar*	Quench	*Thirst*	Touchdown	*Football*
Fracture	*Break*	Plates	*Cups*	Stubborn	*Mule*
Delicate	*Fragile*	Caffeine	*Coke*	Pail	*Bucket*
Promise	*Broken*	Wrench	*Screwdriver*	Spank	*Slap*
Cork	*Bottle*	Hut	*Straw*	Deer	*Buck*
Sweet		**Color**		**Cotton**	
Bee	*Honey*	Marker	*Crayon*	Slacks	*Polyester*
Cane	*Sugar*	Reading	*Rainbow*	Material	*Fabric*
Lollipop	*Candy*	Drag	*Race*	Powder	*Puff*
Adorable	*Cute*	Plum	*Purple*	Sheep	*Wool*
Fudge	*Chocolate*	Curtain	*Shade*	Plush	*Fluffy*
Polite	*Nice*	Clear	*Vivid*	Peach	*Fuzz*
Pancakes	*Syrup*	Art	*Paint*	Satin	*Silk*
Sincerity	*Kindness*	Preference	*Favorite*	Rag	*Cloth*
Dance		**Cheese**		**Snow**	
Graceful	*Ballet*	Noodles	*Macaroni*	Slope	*Ski*
Tuxedo	*Prom*	Chili	*Crackers*	Reindeer	*Sleigh*
Pat	*Tap*	Hut	*Cottage*	Corn	*Flake*
Blues	*Jazz*	Cauliflower	*Broccoli*	Coat	*Frost*
Hoop	*Hula*	Pepperoni	*Pizza*	Dig	*Winter*

Table 5 (continued)

Hymn	*Song*	Turkey	*Ham*	Umbrella	*Shovel*
Member	*Club*	Rat	*Mouse*	Thaw	*Rainbow*
Act	*Perform*	Grill	*Hamburger*	Climber	*Melt*
Finger		**Hair**		**Dark**	
Glove	*Hand*	Fingernail	*Clippers*	Bulb	*Light*
Hitchhike	*Thumb*	Itch	*Lice*	Bat	*Cave*
Brass	*Knuckle*	Honey	*Comb*	Underground	*Tunnel*
Feet	*Toes*	Flex	*Shampoo*	Noon	*Midnight*
Hammer	*Nail*	Bandana	*Headband*	Sun	*Tan*
Unit	*Digit*	Flakes	*Dandruff*	Fade	*Dim*
Movie	*Flick*	Foam	*Mousse*	Curtain	*Shade*
Tweezers	*Splinter*	Eagle	*Bald*	White	*Black*
Grape		**Clock**		**Hot**	
Jungle	*Vine*	Snooze	*Alarm*	Chill	*Cold*
Nuts	*Raisin*	Moment	*Time*	Thermometer	*Temperature*
Bitter	*Sour*	Snap	*Click*	Refreshing	*Mile*
Beetle	*Juice*	Flea	*Tick*	Blaze	*Cool*
Violet	*Purple*	Observe	*Watch*	Range	*Fire*
Peanut Butter	*Jelly*	Counter	*Clockwise*	Blanket	*Stove*
Seasoning	*Flavor*	Doorbell	*Chime*	Radiator	*Warm*
Traffic	*Jam*	Curved	*Winding*	Toaster	*Heat*
Slick		**Feather**		**Island**	
Clever	*Sly*	Asphalt	*Tar*	Explorer	*Adventure*
Suds	*Soap*	Geese	*Birds*	Hula	*Hawaii*
Secretive	*Sneaky*	Tribe	*Indian*	Pearl	*Harbor*
Petroleum	*Vaseline*	Buffalo	*Wings*	Climate	*Tropical*
Daring	*Cunning*	Hard	*Soft*	Chest	*Treasure*
Snail	*Slimy*	Cushion	*Pillow*	Secluded	*Isolated*
Seal	*Otter*	Fried	*Chicken*	Outlet	*Inlet*
Seller	*Salesman*	Pigeon	*Dove*	Shack	*Hut*

Critical items are listed in bold; nonpresented mediators are listed in italics.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261–295. doi:10.1016/S0022-5371(83)90201-3

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123. doi:10.1037/h0033773

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: V. H. Winston.

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 12*, 336–345. doi:10.1037/0278-7393.12.3.336

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftus, B., et al. (2007). The english lexicon project: A user's guide. *Behavior Research Methods, 39*, 445–459.

Bäulm, K. H., & Kuhbandner, C. (2003). Retrieval-induced forgetting and part-list cuing in associatively structured lists. *Memory & Cognition, 31*, 1188–1197.

Bousfield, W. A. (1953). The occurrence of clustering in recall of randomly arranged associates. *The Journal of General Psychology, 49*, 229–240.

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*, 3–47. doi:10.1016/0273-2297(90)90003-M

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164–169. doi:10.1111/1467-8721.00192

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in "associative" false memory. *Psychonomic Bulletin & Review, 15*, 1035–1053. doi:10.3758/PBR.15.6.1035

Chwilla, D. J., & Kolk, H. H. J. (2002). Three-step priming in lexical decision. *Memory & Cognition, 30*, 217–225.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407–428. doi:10.1016/0010-0277(92)90041-F

Cortese, M. J., Khanna, M. M., & Hacker, S. (2010). Recognition memory for 2, 578 monosyllabic words. *Memory, 18*, 595–609. doi:10.1080/09658211.2010.493892

Cotel, S. C., Gallo, D. A., & Seamon, J. G. (2008). Evidence that nonconscious processes are sufficient to produce false memories. *Consciousness and Cognition, 17*, 210–218. doi:10.1016/j.concog.2007.01.009

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22. doi:10.1037/h0046671

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

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38*, 833–848. doi:10.3758/MC.38.7.833

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

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative activation theory of children's and adult's memory illusions. *Journal of Memory and Language, 60*, 229–251. doi:10.1016/j.jml.2008.10.002

Huff, M. J., Meade, M. L., & Hutchison, K. A. (2011). *Age-related differences in guessing on free and forced recall tests* (under review)

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20*, 497–514. doi:10.1016/S0022-5371(81)90138-9

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A *micro*-analytic review. *Psychonomic Bulletin & Review, 10*, 785–813.

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. doi:10.1016/j.jml.2004.08.003

Hutchison, K. A., & Davis, S. D. (2010, November). *Testing mediated priming using ambiguous and unambiguous mediators: Will "kidney" prime "piano"?* Poster presented at the 51st Annual Meeting of the Psychonomic Society, St. Louis, MO.

Jones, L. L. (2010). Pure mediated priming: A retrospective semantic matching model. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*, 135–146. doi:10.1037/a0017517

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavioral Research Methods, Instruments, & Computers, 28*, 203–208.

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

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27*, 545–559. doi:10.1016/0749-596X(88)90025-3

Meade, M. L., Watson, J. M., Balota, D. A., & Roediger, H. L., III. (2007). The roles of spreading activation and retrieval mode in producing false recognition in the DRM paradigm. *Journal of Memory and Language, 56*, 305–320. doi:10.1016/j.jml.2006.07.007

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234. doi:10.1037/h0031564

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97. doi:10.1037/h0043158

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254. doi:10.1037/0096-3445.106.3.226

Neely, J. H., Johnson, J. D., Neill, W. T., & Hutchison, K. A. (1999, November). *Is there memory to be found in "false memory"?* Presented at the 40th Annual Meeting of the Psychonomic Society, Los Angeles, CA.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. (1999). The University of Florida word association, rhyme and word fragment norms. Available from http://w3.usf.edu/FreeAssociation

Rabinowitz, J. C., Craik, F. I. M., & Ackerman, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology, 39*, 325–344. doi:10.1037/h0080643

Robinson, K. J., & Roediger, H. L., III. (1997). Associative processes in false recall and false recognition. *Psychological Science, 8*, 231–237. doi:10.1111/j.1467-9280.1997.tb00417.x

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. doi:10.1037/0278-7393.21.4.803

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 & Review, 8*, 385–407.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., et al. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition, 30*, 1054–1064.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale: Erlbaum.

Tse, C.-S., & Neely, J. H. (2005). Assessing activation without source monitoring in the DRM false memory paradigm. *Journal of Memory and Language, 53*, 532–550. doi:10.1016/j.jml.2005.07.001

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

Author Note