Original Article

QJEP

Quarterly Journal of Experimental Psychology
2021, Vol. 74(3) 483–496

DOI: 10.1177/1747021820965061
qjep.sagepub.com
SAGE

Mapping the time course of semantic activation in mediated false memory: Immediate classification, naming, and recognition

Mark J Huff[1], Alyssa Di Mauro[1], Jennifer H Coane[2] and Liam M O'Brien[3]

Abstract

We evaluated the time course of persistent automatic spreading activation from a mediated list of indirect associates (e.g., *meow, day*, and *basement*) that all converged upon a non-presented critical item (CI; e.g., *black*). Mediated lists were related to CIs through non-presented mediators (e.g., *cat, night*, and *bottom*). Three speeded tasks were used to evaluate the time course of semantic activation of the CI: a continuous semantic classification task (concrete/abstract decisions), a naming task (reading words aloud), or a recognition test (old/new memory decisions). Test lists were presented immediately following the mediated lists, and CIs were presented in the first, third, or eighth positions. The results revealed that in both the classification and naming tasks, CI priming was greatest in the first test position and declined across the remaining test positions. Importantly, priming was statistically reliable in the late test positions, providing evidence for long-term semantic priming (i.e., across positions on immediate tasks). False recognition, however, was stable across test positions. Collectively, these patterns suggest that spreading-activation processes decline, consistent with implicit spreading activation, and these processes may contribute to long-term false recognition.

Keywords

False memory; mediated false memory; recognition; semantic classification; speeded naming

Received: 28 February 2020; revised: 1 September 2020; accepted: 6 September 2020

The effects of associative or semantic relations across memory sets have been well-documented by memory researchers. Enhancing the relatedness across items often aids memory as shown by improved memory for related versus unrelated lists and word pairs (Castel et al., 2007; Koriat & Bjork, 2005; Rabinowitz et al., 1982), a tendency to generate clusters of items based on category membership (Bousfield, 1953; Huff & Bodner, 2014; Miller, 1956), and improved memory when elaborative encoding tasks emphasise relational processing over a read-only control (Huff & Bodner, 2019).

Although associative processing or relatedness in general can yield memory benefits, it can also yield memory costs. A common demonstration of costs resulting from associative processing is the Deese/Roediger–McDermott illusion (DRM; Deese, 1959; Roediger & McDermott, 1995), in which participants study lists of associates (e.g., *bed, rest, tired*, and *dream*) that all converge upon a single, non-presented critical item (CI; *sleep*). According to activation models of memory (Anderson, 1983; Collins & Loftus, 1975), processing any item results in an automatic spread of activation to its associates, thereby facilitating lexical access to the latter. Original DRM lists were generated using word association norms such that list items consisted of the most common responses when cued with the CI. These lists are, therefore, strongly related both associatively and thematically, which is why false recall and false recognition of the CI are robust (Gallo, 2010, for review; Lampinen et al., 1999; Roediger & McDermott, 1995), are often resilient to various methods designed to reduce the

[1]School of Psychology, The University of Southern Mississippi, Hattiesburg, MS, USA
[2]Department of Psychology, Colby College, Waterville, ME, USA
[3]Department of Mathematics and Statistics, Colby College, Waterville, ME, USA

Corresponding author:
Mark J Huff, School of Psychology, The University of Southern Mississippi, Hattiesburg, MS 39406, USA.
Email: mark.huff@usm.edu

illusion (Gallo et al., 2001; Huff & Bodner, 2013; McCabe & Smith, 2002), and persist on memory tests completed after a significant delay (Seamon et al., 2002). Given the persistence of list-based false memory illusions, particularly after a delay, the purpose of the present study is to examine the persistence of long-term CI activation following exposure of several unrelated items using a series of speeded response tasks.

Thematic and associative processes in the DRM illusion

Although several mechanisms have been proposed to account for the DRM illusion (see Gallo, 2006, for review), the two most frequently cited are fuzzy-trace theory (FTT; Brainerd & Reyna, 1990) and activation monitoring theory (AMT; Roediger, Balota, & Watson, 2001). FTT posits that during encoding of DRM lists, participants store both a verbatim and gist memory representation (Brainerd & Reyna, 2002). The verbatim representation corresponds to specific details of list items themselves such as the font, colour, and test position. The gist representation contains the thematic meaning of the item or group of items when meanings are consistent in the list. Because DRM CIs are not presented in the list, the illusion is due to the *extraction* of thematic gist information, which is incorrectly endorsed as studied during the test phase. The DRM illusion can be mitigated, however, provided recollection of sufficient verbatim information allows for rejection of gist information (i.e., recollection rejection). In addition, gist representations are suggested to be relatively stable over time, whereas verbatim representations generally decay or become interfered with, leading to a reduction in memory for studied list items. Consistent with the predictions of FTT, the DRM illusion has been shown to persist over a delay of weeks or even months (Seamon et al., 2002), despite memory for studied list items declining.

In contrast, AMT involves a two-stage process in which the DRM illusion is due to automatic lexical spreading activation from studied list words to the CI, and a monitoring failure at test in which individuals falsely endorse CIs as presented on the initial study list (presumably by misattributing the increased familiarity or accessibility of the CI to a study event instead of implicit activation processes). Associative-based processes, such as lexical spreading activation, have also received empirical support. Deese (1959) showed that backward associative strength (BAS)—a measure of association between studied list items and the CI—was related to the likelihood that CIs would be falsely recalled. Similarly, Roediger, Watson, et al. (2001), using a regression analysis, reported that BAS was the strongest predictor of CI false recall and the second strongest predictor of CI false recognition. Consistent with this pattern, longer DRM lists, which are likely greater in total BAS, are positively related to the DRM illusion (e.g., Coane et al., 2007; Hutchison & Balota, 2005; Robinson & Roediger, 1997), providing additional evidence that implicit activation of the CI is related to the DRM illusion. However, it is worth noting that thematic coherence and specific types of semantic features not only relate to the DRM illusion but are also related to BAS (Cann et al., 2011). Collectively then, in the DRM paradigm, gist extraction and activation monitoring are perfectly confounded—study lists possess a strong thematic consistency supporting gist extraction, but also share strong associations to the CI, supporting spreading activation. Furthermore, both theoretical accounts include an error-increasing mechanism (gist extraction and automatic activation) and an error-reducing mechanism (verbatim recollection and test-based monitoring). As such, both accounts can accommodate the DRM findings reported in the literature.

To parse the thematic/activation accounts of the DRM illusion, researchers have had to rely upon newly constructed lists that minimise the contributions of gist processes leaving associative processes intact. Hutchison and Balota (2005) compared false memories on homograph lists relative to traditional DRM lists that were matched on BAS. On homograph lists, list items were related to a CI (e.g., *right*), but the list contained items from two separate meanings, which disrupted the consistency of the gist representation (e.g., Meaning 1: *wrong, correct, accurate*, etc.; Meaning 2: *left, handed, starboard*, etc.). Despite the presence of separate themes, homograph and DRM lists produced equivalent false recall/recognition, supporting activation-monitoring processes over gist extraction. Separately, however, Huff et al. (2015) found evidence for gist extraction using homograph lists, but only when gist processes were encouraged by blocking words together by homograph theme and using a delayed test. Thus, homograph lists did not consistently support either account.

More recently, stronger evidence supporting AMT has been reported by Huff and colleagues (Coane et al., 2016; Huff et al., 2012; Huff & Hutchison, 2011), using mediated lists in which list items are only indirectly related to CIs through non-presented mediators. For instance, studying list items *faucet, London, jog*, and so on can produce false recognition of the CI *river*, through implicit activation of a set of mediators that are not studied (e.g., *water, bridge, run*). This *mediated false recognition* pattern is noteworthy because mediated list items are completely unrelated to each other and are not directly related to the CI according to the associative norms. The unrelated nature of mediated lists precludes the formation of a consistent gist representation and therefore could only be due to spreading activation processes. Although the evidence for activation-monitoring theory in associative false memory using mediated lists is quite convincing, we note several recent instances that attribute the DRM illusion as occurring through gist-based processes (Abadie & Camos, 2019; Cann et al., 2011; Reyna et al., 2016), indicating that the debate remains unresolved. Although specifying those theories which explain and predict factors that contribute to memory illusions are critical for applied considerations, the aim of the present experiment was to evaluate implicit activation of mediated critical lures via a semantic priming paradigm, rather than adjudicating between theoretical explanations provided by AMT and FTT directly.

The semantic priming approach to examining activation-based processes

The major assumption of previous studies on mediated false memories is that the presentation of study lists implicitly activates mediators, which in turn, activate the CI through spreading-activation processes. The goal of the present study is to provide a more direct assessment of activation processes of mediated study lists using a semantic priming paradigm—a common method used to gauge implicit lexical processing. In this paradigm, participants are required to make a series of decisions or responses to a series of words. Common tasks include the lexical-decision task (LDT; word vs. non-word decisions), the semantic classification task (determine whether a word is concrete vs. abstract), and the pronunciation/naming task (reading words aloud). These tasks have often yielded a semantic priming effect: faster responses to a target word (e.g., *window*) when preceded by a related prime (e.g., *blinds*) versus a non-related prime (e.g., *pencil*; Meyer & Schvaneveldt, 1971; Neely, 1977; see McNamara, 2005, for review). Semantic priming has been interpreted as evidence for spreading-activation processes, especially when post-lexical strategies, such as semantic matching, have been minimised when tasks are completed continuously (McNamara & Altarriba, 1988).[1]

Spreading activation processes have also been supported using prime–target pairs that are associated through non-presented mediators. For instance, facilitation is found when target words (e.g., *box*) are preceded by prime words (e.g., *beach*) that are related by a non-presented mediator (e.g., *sand*), compared to unrelated pairs (Balota & Lorch, 1986; Chwilla & Kolk, 2002; Hutchison, 2003; Jones, 2010). Mediated primes and targets typically yield semantic priming effects that approximate half the magnitude of directly related items (Balota & Lorch). Indeed, mediated priming experiments served as inspiration for investigating associative false memory effects reported above (e.g., Huff et al., 2012; Huff & Hutchison, 2011).

Mediated priming is consistent with an automatic spreading activation account as it suggests a two-step process in which activation from primed concepts spreads to adjacent concepts that are directly and indirectly related, facilitating downstream responses (Anderson, 1983; Collins & Loftus, 1975; Hutchison, 2003). Some researchers (e.g., Chwilla et al., 2000; Hill et al., 2002; Jones, 2010, 2012) have suggested that a retrospective semantic matching process could account for some mediated priming effects, particularly when associations between prime, mediator, and target are weak. However, mediated priming effects have been reliably found in pronunciation tasks, in which strategic semantic matching processes are generally thought to be minimised (Balota & Chumbley, 1984; Jones & Estes, 2012), providing support for spreading activation in mediated priming.

In the DRM paradigm, spreading activation as an underlying mechanism hinges, in part, upon whether activation generated at encoding can persist over a delay or decays over time. The DRM illusion has been found to persist for weeks to months following study (Seamon et al., 2002; Thapar & McDermott, 2001; but see Colbert & McBride, 2007), though semantic priming effects between a single prime and target are generally short-lived (typically lasting 1–2 s) and can be reduced or eliminated with a single intervening item (Dannenbring & Briand, 1982; Masson, 1995) or when participants attend to divergent semantic information (Balota et al., 1992; Neely, 1977). The persistence of false memory effects has often been used as a line of evidence against activation and in support of gist extraction as a driving mechanism in the DRM illusion, because, it is argued that any activation effects would have dissipated long before a delayed test is administered (e.g., Brainerd et al., 2008). However, Joordens and Becker (1997; see also Ray & Bly, 2007) have reported long-term semantic priming effects across several intervening items when response tasks encouraged processing of semantic features. Clearly, in this context, "long-term" refers to *priming* effects persisting beyond 1–2 s and several intervening items, not to persistent memory traces. Because DRM lists consist of many semantically related "primes," activation from DRM lists could persist beyond that of a single prime (see Tse & Neely, 2005, 2007, for review). Thus, given the apparent inconsistency in the literature between the persistence of false memory effects and the short-lived nature of activation, further elucidating the time course and duration of activation processes is critical for explaining the phenomenon and its potential generalisability to other memory illusions.

Long-term priming of DRM CIs

Several studies have examined long-term semantic activation originating from DRM materials. Zeelenberg and Pecher (2002) and McKone (2004) had participants initially study many DRM lists followed by an LDT on a set of items which included CIs, DRM list words, and nonwords. In both studies, no evidence of CI priming was found, which may reflect decay of activation due either to the presence of several intervening lists or a relatively long delay between study and the LDT (cf. Tse & Neely, 2005). In contrast, studies that employ an LDT immediately after studying each DRM list have found evidence for CI priming (e.g., Hancock et al., 2003; Meade et al., 2007; Tse & Neely, 2005, 2007); suggesting that the LDT must be completed shortly after the presentation of the DRM list to detect reliable priming. However, differences in the persistence of priming effects over time emerged across studies, with some showing priming across different test positions, including long-term priming at later positions (e.g., when the CI was the third or eighth item on the test list), and others only showing priming when the CI was in the first and most immediate test position.

In one example, Tse and Neely (2005) presented participants with a series of 14-item DRM lists without non-words. Each list was followed by a 30-s filler task and a 40-item LDT that presented the CI in either the first or second half of the LDT trials. Relative to a baseline list, in which CIs were presented following the presentation of an unrelated list, Tse and Neely found evidence for long-term semantic priming when the CI was presented on both the first and second halves of the LDT. Priming in the second half was diminished in two of their four experiments—a pattern consistent with a decay of activation characteristic of spreading activation processes. The long-term priming pattern was replicated in a subsequent study (Tse & Neely, 2007) using conditions that controlled for test-induced priming by withholding semantically related test items prior to the presentation of the CI (cf. Coane & McBride, 2006).

In a separate study, Meade et al. (2007) presented 27-item lists that contained 15 DRM items that were arranged such that blocks of five DRM items were interleaved with blocks of four items that were either unrelated words or non-words. After a brief 1-s delay, participants completed an LDT in which CIs were presented in test Positions 1, 3, 6, or 11, to gauge the time course of semantic priming. Reliable CI priming was found, but only in the first test position. Priming decayed and was not reliable by the third test position. In a separate condition, participants were presented with the same 27-item lists, but instead completed a speeded recognition test on the same items. In contrast to the LDT results, false recognition rates were reliable and stable across test positions, demonstrating a disconnect between implicit CI activation as assessed on the LDT and false recognition. To account for these differences, Meade et al. suggested that recognition test instructions may encourage the use of a retrieval-mode process (Tulving, 1983), in which accessing episodic memory may cause reactivation of the associative network established during encoding, contributing to long-term false recognition.

In another study, Meade et al. (2010) further examined long-term CI priming by evaluating potential additivity of priming from DRM lists by presenting participants with 14-item DRM lists in which either the first 7 items were related DRM items and the last 7 were unrelated, the first 7 were unrelated and the last 7 were related, or all list items were related (cf. Balota & Paul, 1996). Each list was then immediately followed by a naming task in which participants read aloud each word and vocal onset responses were measured. The shift to a naming task was done to eliminate non-words from the study list that were used for the LDT and to reduce potential contributions of post-lexical decision strategies (Neely, 1991). The CI was presented in test Position 1, 3, or 9 to gauge long-term priming. Importantly, long-term priming was found in later test positions, particularly when the first seven items were strongly related and the last seven were unrelated. Thus, although there are some inconsistencies regarding long-term priming, the general pattern across studies seems to favour the presence of long-term CI priming with some evidence that priming decays over time.

Present study

In this study, we further evaluate long-term semantic priming effects of CIs by employing indirectly related mediated lists. The use of mediated lists to examine long-term priming of CIs is advantageous over directly related DRM lists for two reasons. First, activation of CIs is implicit in nature as the study list items are not related to CIs based on the Nelson et al. (1999) associative norms and participants are generally unaware of the association between the study list and the CI (Coane et al., 2016; Huff et al., 2012). Long-term priming effects would, therefore, only reflect implicit associations between study lists. Second, the use of mediated lists greatly restricts, and possibly even prevents, the use of post-lexical response strategies which would operate outside of implicit activation. This is because even if participants were to employ a post-lexical strategy when responding to test items, this strategy would not be beneficial as test items share no direct association with the studied mediated list and would not be explicitly predictive. Indeed, Meade et al. (2010) suggested that the greater long-term CI priming found when the first seven items studied were strongly related and the last seven unrelated may have been due to participants explicitly generating a theme label for these items when they noticed a difference with the second half being unrelated items. They then relied upon this label during the naming task which produced long-term priming. The label would presumably be the theme or gist of the list, consistent with the notion that gist extraction occurs rapidly. This process would be unlikely to occur with mediated lists given the indirect associations preclude conscious identification of a theme label (cf. Coane et al.; Huff et al.). Thus, the use of mediated lists to evaluate long-term priming of CIs is a stronger test of implicit CI activation by controlling for thematic contributions present in DRM lists and further limits post-lexical response strategies such as explicit label generation.

Building off previous long-term DRM priming studies, our study had three aims. First, we sought to establish whether study of mediated lists would produce long-term semantic priming effects. To this end, our methodology was inspired by Meade et al. (2007, 2010), in which CIs were embedded in test lists at different test positions to examine the time course of mediated priming. We, therefore, compared response latencies to CIs when they were presented in test Position 1, 3, or 8, relative to control items presented at the same positions. These comparisons allowed us to evaluate whether priming effects occur long-term (i.e., priming on test Position 8) or only short-term (i.e., priming only on test Position 1). Based on DRM studies reviewed above (e.g., Hancock et al., 2003; Meade et al., 2010; Tse & Neely, 2005, 2007), we anticipated that priming would occur on later test positions, consistent with long-term priming.

Our second aim was to evaluate whether long-term priming effects were sensitive to different speeded response tasks at test. Semantic contributions in visual word recognition paradigms tend to be stronger using tasks that emphasise the processing of semantic features than tasks that do not. For instance, Yap et al. (2012) found that semantic richness variables are stronger predictors of response latencies on a speeded semantic classification task (e.g., concrete/abstract response decisions) than a naming task where items are simply read aloud. Given these task differences, detection of long-term priming effects may be contingent upon the semantic processing fostered by the task completed at test. In addition, relative to directly related items, priming from mediated items tends to be relatively modest (see Hutchison, 2003), which may necessitate response tasks that are more sensitive to semantic effects. We therefore tested task effects on mediated priming by comparing latencies on a semantic classification task to those on a naming task. Finally, our third aim was to evaluate the long-term time course of mediated false recognition across test positions. We expected that mediated false recognition would persist across test positions consistent with Meade et al. (2007, 2010) using DRM lists.

In sum, we measured the time course of mediated CI priming and false recognition using three between-subject groups. Following study of mediated lists, participants either completed a semantic classification task (SCT), a speeded naming task, or an old/new recognition test (RGN). To assess priming and false recognition patterns, task/test lists that contained CIs were compared with filler lists that contained unrelated control items that were matched to CIs based on length, frequency, and concreteness to control for potential item effects.

Method

Participants

A total of 427 undergraduates participated for partial course credit. Of these participants, 264 were tested at The University of Southern Mississippi and 163 at Colby College. All participants reported proficiency in the English language and had normal or corrected-to-normal vision. Participants were randomly assigned to SCT, naming, or RGN groups. Twenty-six were removed from analyses due to either repeated failures to respond under the 1,000 ms deadline resulting in excessive timeouts (15% or more of the test trials; $n = 13$), due to a computer error ($n = 5$), or randomly tapping keys during the test lists ($n = 8$), leaving 136 in the SCT group, 129 in the naming group, and 136 in the RGN group. Removed participants were equally distributed across testing sites, leaving a total of 401 for analysis. A sensitivity power analysis computed using *G*Power* (Faul et al., 2007) indicated that the sample size had sufficient power (.80) to detect small effect sizes (Cohen's $d = 0.20$) or larger, which is appropriate given mediated priming effects are generally small (Hutchison, 2003).[2]

Materials

Twenty-four mediated lists were taken from Huff et al. (2012) and used as study lists in the experiment. Mediated lists consisted of 15 indirectly related words (e.g., *meow, day*, and *basement*) that converged upon a non-related CI (e.g., *black*) through non-presented mediators (e.g., *cat, night, bottom*; see Table 1 for list construction). The non-presented mediators were items from original DRM lists. Each mediated list was then combined with four unrelated buffer words for a total of 19 words. Buffers were included to provide a set of correct items to test for in the RGN group and to examine repetition priming in the SCT and naming groups. Buffers were presented in the first four study positions to ensure that mediated list items were grouped together and were the most recent items studied prior to the test list. Study lists were presented in the same order and the lists were identical across all three groups.

Test lists for each of the 24 lists were also identical across groups. Each test list contained 10 items: Three previously studied buffers and seven unstudied items in which one of these items was either the CI or critical control item and the other six were unrelated distractors. Half of the test lists contained a CI related to the studied mediated list, whereas the other half contained a critical control item. Unstudied items were matched to the studied mediated list items using the English Lexicon Project database (Balota et al., 2007) based on word length, word frequency (using the SUBTL database; Brysbaert & New, 2009), and concreteness (MRC Psycholinguistic database; Coltheart, 1981). Mediated CIs were further equated to a set of critical control items using the same factors to provide an appropriate comparison to compute priming effects. Importantly, on the test lists, the CI was the only item related to the mediated lists to eliminate the possibility for test-induced priming (Coane & McBride, 2006; Marsh et al., 2004). Test lists were constructed such that critical

Table 1. Example 19-item mediated study list (non-presented DRM list in parentheses) and the test lists for critical lures presented in Positions 1, 3, and 8 for the recognition, semantic categorisation, and naming groups.

Study list items	Test lists		
Critical Lure: **Black**	Lure Position 1	Lure Position 3	Lure Position 8
nation			
microscope			
career			
Capital			
meow (cat)	**black**	*nation*	*nation*
day (night)	*nation*	partner	partner
basement (bottom)	partner	**black**	envelope
hue (colour)	envelope	envelope	forbid
red (blue)	forbid	forbid	salad
life (death)	salad	salad	*career*
pale (white)	*career*	*career*	tribe
shadow (dark)	tribe	tribe	**black**
burial (funeral)	payment	payment	payment
pen (ink)	*capital*	*capital*	*capital*
master (slave)			
tan (brown)			
matter (grey)			
mine (coal)			
sorrow (grief)			

Bolded items denote non-presented critical lures, and italicised items reflected unrelated studied buffers.

and critical control items were always presented in the first, third, or eighth test positions. CIs and critical controls in each of the three positions for each list were counterbalanced across participants. For the remaining test positions, buffers or unrelated unstudied items were once randomised and presented in the same order for all participants. Each participant saw every list and each list was presented in each of the experimental conditions an approximate equal number of times across participants. The order of the 24 study lists and their corresponding test list were once randomised and presented in two blocks of 12 lists to allow for a mid-experiment participant break. Block orders were similarly counterbalanced. Thus, the test position order was not blocked in any systematic order (cf. Meade et al., 2007, Experiment 1). Table 1 provides an example study and test list with CI orderings.

Procedure

Participants were tested individually on a computer using E-prime 3.0 software (Psychology Software Tools, 2016) and responses were made using a keyboard in the SCT and RGN conditions and a microphone relayed to an external response box in the naming group. Participants were instructed that they would be presented with 24 pairs of lists in which for each pair they would silently read the first list (the study phase) and then complete a speeded memory test on the second list (the test phase). They were instructed to pay close attention to each of the words during the study phase given the upcoming test. Note that the SCT and naming groups, participants were informed of an upcoming test even though a test was not completed. In the RGN group, participants determined whether the words were presented on the first list by making old/new recognition response. In the SCT group, the speeded task required participants to make concrete/abstract judgements for each word on the second list. In the naming group, participants read each word on the second list aloud into a microphone, which measured the onset of their vocal response. Participants were given a practice set which consisted of two unrelated lists so they could practice their task prior to beginning the experiment. Task instructions were presented at the beginning of the experiment and repeated after the practice list to ensure that participants understood the task.

During the study phase, items were presented at the centre of the computer screen for 1,000 ms with a 500 ms inter-stimulus interval. After the presentation of each list, a loud tone was presented for 1,000 ms, which indicated the start of the test phase. In the test phase, each test item was presented for a maximum of 1,000 ms and participants were required to make their task-specific response within this timeframe or else the trial would time out and participants would receive feedback encouraging them to respond faster, which was displayed for 1,000 ms. Participants were informed of the response deadline prior to beginning the experiment and were encouraged to either rest their index fingers on the labelled keyboard keys to make their condition-specific responses quickly (in the SCT and RGN groups), or to place their mouth close to the microphone so they were ready to read the test phase words (in the naming group). After each test phase, participants saw a screen that said "next list" for 1,500 ms to cue them for the upcoming study phase. This procedure repeated for 12 study/test cycles followed by a participant-paced break and then an additional 12 study/test cycles. At the completion of the final test phase, participants provided demographic information and completed the Shipley vocabulary test, the F-A-S verbal fluency and category fluency tests (Spreen & Benton, 1977), and a verbal debriefing. Neither the vocabulary nor the fluency tests were found to be related to the priming effects and therefore are not discussed further. A typical experimental session lasted approximately 60 min.

Results

Our analyses assess the time course of semantic activation of mediated critical lures across test positions on the

Table 2. Mean (95% CI) proportions of false alarms and mean reaction times (RTs, in microseconds) to critical items and critical item controls as a function of test position for the semantic categorisation, naming, and speeded recognition groups.

Group/Test item	Test position		
	Position 1	Position 3	Position 8
Semantic categorisation group (*N* = 136)			
Critical item RTs			
Control critical items	684 (16)	616 (13)	638 (16)
Critical items	646 (15)	602 (14)	618 (16)
Priming effect	38	14	20
Naming group (*N* = 129)			
Critical item RTs			
Control critical items	561 (13)	509 (12)	510 (11)
Critical items	541 (14)	498 (11)	494 (10)
Priming effect	20	11	16
Recognition group (*N* = 136)			
Critical item false recognition			
Control critical items	0.27 (0.05)	0.15 (0.04)	0.17 (0.04)
Critical items	0.32 (0.05)	0.20 (0.04)	0.23 (0.04)
False recognition effect	0.05	0.05	0.06
False alarm RTs			
Control critical items	680 (34)	625 (36)	609 (37)
Critical items	690 (31)	628 (34)	623 (31)
Difference	10	3	14
Correct rejection RTs			
Control critical items	659 (18)	600 (13)	592 (13)
Critical items	672 (19)	610 (15)	609 (16)
Difference	13	10	17

SCT and naming tasks and compare this activation with false recognition found across positions on the RGN task. To foreshadow our results, semantic priming persisted across test positions in both the SCT and naming tasks, though priming was greater in the first test position. False recognition persisted equivalently across test positions. Linear mixed models were completed for all response latency and recognition analyses reported below. All analyses were completed in *R* 3.5.1 (www.r-project.org). Linear mixed-effects model analyses were conducted using the *lmer* function and generalised linear mixed-effects models were run using the *glmer* function. Note that model analyses examining position effects were conducted by comparing responses to test Position 1, relative to Positions 3 and 8. Thus, the tables which include the position output from the mixed models include rows pertaining to Positions 3 and 8, which indicate a comparison to Position 1. In addition, Bayes factors (BFs) were included when comparing models and were computed using the Bayesian information criterion (BIC) fit statistics provided by the *lmer* and *glmer* functions using the *BayestestT* package. BF values assess the evidence for the more complicated (larger) model compared with the less complicated (smaller) model. We use the convention established by Kass and Raftery (1995), barely worth mentioning (1–3.2), substantial (3.2–10), strong (10–100), and decisive (>100), in which BFs greater than 1 provide evidence for the alternative hypothesis.

Semantic classification and naming tasks

Response latencies. Response latencies made between 350 and 1,000 ms and within 2.5 *SD*s of the mean for each participant were included in the analysis. In the SCT group, the trimming procedure removed 7.6% of Position 1 trials, 3.4% of Position 3 trials, and 3.9% of Position 8 trials. In the naming group, trimming removed 4.9% of Position 1 trials, 3.3% of Position 3 trials, and 2.1% of Position 8 trials. No participant had more than two trials removed per cell out of a maximum of four due to the trimming procedure and most participants had all four trials available in each cell. Trials removed were evenly distributed across trial types. For the analyses reported, we collapsed across concrete and abstract responses in the SCT given no a priori reason to expect a difference between response types.

Mean reaction times for CIs and critical controls across three possible test positions for the SCT and naming groups are displayed in Table 2 (along with the difference scores). Response latencies were modelled via a linear

Table 3. Summary of linear mixed-model analyses for reaction times for critical items and critical item controls as a function of test position for the semantic categorisation and naming task groups.

Statistic/comparison	Estimate	*SE*	*t*	*p* value	Omnibus χ^2	*p* value
Intercept	674.3	8.38	80.4	$<1.0\times10^{-6}$	–	–
Group = naming	115.4	8.41	13.7	$<1.0\times10^{-6}$	–	–
Position = 3	50.2	5.81	8.6	$<1.0\times10^{-6}$	87.0	$<2.0\times10^{-16}$
Position = 8	42.5	5.76	7.4	$<1.0\times10^{-6}$		
Item type = CI	20.0	9.04	2.21	.027	–	–
Pos. = 3 × item = CI	4.4	6.91	0.6	.524	4.8	.089
Pos. = 8 × item = CI	10.3	6.76	1.52	.129		
Group = naming × Pos. = 3	8.6	6.37	1.35	.177	3.32	.016
Group = naming × Pos. = 8	9.6	6.38	1.50	.133		

SE: standard error; CI: critical item; SCT: semantic classification task; BF: Bayes factor.
For test position comparisons, Positions 3 and 8 in the table reflect model comparisons to Position 1. For the group comparison, naming in the table reflects a model comparison to the SCT. For the item type comparison, CI in the table reflects a model comparison to CI controls. BF > 100 in support of this model against an intercept-only model with random effects.

mixed-modelling approach. The initial model included main effects for task type (naming vs. SCT), item type (CI vs. critical control), position (1 vs. 3 vs. 8), and the interactions. Random offsets were allowed for each participant, as well as for each test word. Likelihood ratio tests (LRTs) confirmed the necessity of including both random effects in the model ($\chi_1^2 = 240.4$, $p = 9.28 \times 10^{-61}$; BF > 100). Model results and the BF are reported in Table 3. The final model included significant fixed effects for the main effect of task type, which indicated that naming latencies were faster than SCT latencies (519 vs. 634 ms), the main effect of item type, which indicated significant priming for CIs over CI controls (567 vs. 586 ms), and the main effect of position, in which latencies were slower for Position 1 relative to Positions 3 and 8 (610 vs. 558 vs. 567 ms, respectively), but the latter two did not differ.

A reliable interaction between task type and test position was also found, which indicated that latencies decreased in naming and SCT tasks from Positions 1 to 3, and increased from Position 3 to 8 in the SCT group (609 vs. 628 ms), but not the naming group (503 vs. 502 ms). Importantly, however, an item type by position interaction was found, which reflected greater CI priming on Position 1 (595 vs. 624 ms, for CIs and CI controls) than Position 3 (551 vs. 564 ms) and Position 8 (558 vs. 576 ms). Importantly, priming remained reliable in Positions 3 and 8, providing evidence for long-term priming of mediated CIs. The priming effect was therefore largest in the first test position and declined across Positions 3 and 8, a pattern consistent with a decay of activation. Neither the task type × item type interaction nor the three-way interaction were reliable in the model.

Repetition priming. We further analysed whether the SCT and naming groups were sensitive to repetition of buffer items that were presented within the study list relative to unstudied items. These items were also compared via linear mixed models (see Table 4 for model results and BF). Fixed effects included main effects for group (SCT vs. naming) and item type (studied buffers vs. non-studied buffers), and the interaction. Necessity of the random offset for test word was confirmed via an LRT ($\chi_1^2 = 240.4$, $p < 1.00 \times 10^{-100}$, BF > 100). From a fixed effects standpoint, only the main effect of group was significant (latencies were faster in the naming vs. SCT; $\chi_1^2 = 9{,}732.7$, $p < 1.00 \times 10^{-100}$, BF > 100), indicating a null repetition priming effect (570 vs. 574 ms for studied buffers and unstudied items, respectively).

Table 4. Summary of linear mixed-model analyses for reaction times studied buffers versus unstudied buffers for the semantic categorisation and naming task groups to test for repetition priming.

Statistic/comparison	Estimate	*SE*	*t*	*p* value
Intercept	674.3	8.38	80.4	$<1.0\times10^{-6}$
Group = naming	125.2	1.21	103.2	$<1.0\times10^{-6}$

SE: standard error; SCT: semantic classification task; BF: Bayes factor.
For the group comparison, naming in the table reflects a model comparison to the SCT. BF > 100 in support of this model versus an intercept-only model.

Speeded recognition

Speeded recognition allowed us to examine whether false recognition effects occurred using identical study and test lists as those used in the SCT and naming tasks above. An important question is whether false recognition patterns would follow the priming patterns found in the SCT and naming tasks. Table 2 reports false alarms to CIs and CI controls as a function of test position. For speeded recognition responses, trials were trimmed using the same criteria as the SCT and naming tasks. This procedure removed 13% of Position 1 trials, and 3% of Position 3 and Position

Table 5. Summary of linear mixed-model analyses for false alarms to critical items and critical control items as a function of test position.

Statistic	Estimate	*SE*	*z*	*p* value	Omnibus χ^2	*p* value
Comparison						
Intercept	1.32	0.16	7.99	$<1.0\times10^{-6}$	–	–
Position 3	0.82	0.13	6.54	$<1.0\times10^{-6}$	73.70	$<2.0\times10^{-6}$
Position 8	0.58	0.12	4.84	$<1.0\times10^{-6}$		
Item type = CI	0.39	0.16	2.39	.017		

SE: standard error; CI: critical item; BF: Bayes factor.
For test position comparisons, Positions 3 and 8 in the table reflect model comparisons to Position 1. For the item type comparison, CI in the table reflects a model comparison to CI controls. BF > 100 in support of this model versus an intercept-only model.

8 trials. Again, no participant had more than two trials removed per cell due to the trimming and most participants had all four trials available in each cell. Removed trials were again distributed evenly across trial types.

False recognition. False alarm rates to CIs and CI controls as a function of test position were similarly modelled using generalised linear mixed models with a logit link function (see Table 5 for model results and BF). Random effects were included both for participants and test words. The need for the random effect of test word in addition to a random effect for subject was confirmed via LRT ($\chi^2_1 = 31.7$, $p = 1.79 \times 10^{-8}$, BF > 100). The model included main effects for item type (CI vs. CI control), position (1 vs. 3 vs. 8), and the interaction. Only the main effects of position and item type were significant, indicating that false alarms were greatest in Position 1 relative to Positions 3 and 8 (.30 vs. 18 vs. 20, respectively), with the latter positions being equivalent, and that false alarms were greater for CIs over CI controls (.25 vs. .20). The interaction was not reliable, indicating that CI false recognition was stable across test positions ($\chi^2_2 = 0.05$, $p = .98$, BF < 0.01).

For completeness, we also analysed response latencies for false alarms and correct rejections for CIs and CI controls (see Meade et al., 2007, for a similar analysis). Note that in this analysis, not every participant had a false alarm or correct rejection for CIs at each test position, leading to missing cells for response latencies. Missing cells were frequent, occurring for 49.1% of the cells across test positions. We therefore collapsed across test position, which decreased the number of missing cells to 21.7%, and note that the cells analysed are relatively unstable due to fewer numbers of available observations. We report mean latencies across test position in Table 2. Latencies for these items were also compared via linear mixed models using random offsets for both participant and test words. LRTs confirmed the need to retain the random offset for test word when considering critical controls ($\chi^2_1 = 16.60$, $p = 4.62 \times 10^{-5}$, BF = 98.60), but did not need to be retained when considering CIs ($\chi^2_1 = 0.68$, $p = .41$, BF = 0.04). Random offsets for subjects were included for both critical control as well as for CIs ($\chi^2_1 = 100.60$, $p = 1.10 \times 10^{-23}$, BF > 100).

Table 6. Mean (95% CI) proportion of hits and reaction times (RTs, in microseconds) for studied buffers and false alarms and RTs to non-studied buffers in the speeded recognition group.

Item type/measure	Studied buffers	Non-studied buffers
"Old" responses	0.69 (0.02)	0.16 (0.03)
RTs	638 (21)	611 (26)

For critical controls, the model predicts the mean response time to be longer for false alarms compared with correct rejections, but only marginally (false alarm mean response time = 627 ms; correct rejection mean response time = 610 ms; $\chi^2_1 = 3.80$, $p = .0513$, BF = 0.16). The same was true for CIs with the mean response time predicted to be longer for false alarms than for correct rejections (false alarm mean response time = 642 ms; correct rejection mean response time = 622 ms; $\chi^2_1 = 6.20$, $p = .01$, BF = 0.60), but this difference was not supported by the BF. Thus, false alarms were only numerically slower than correct rejections, which may be due in part to missing cells and few available observations per participant.

Hits for studied buffers and false alarms for non-studied buffers. Mean hit rates for studied buffer items and false alarms for non-studied buffer items were also compared (see Table 6). Because these items were distributed across the test list, test position was not a factor. Linear mixed models were also used to examine hits for studied buffers and false alarms for unstudied items. Random effects for both test words was found to be necessary in addition to random effects for subjects ($\chi^2_1 = 458.40$, $p = 6.51 \times 10^{-101}$, BF > 100). Item type was found to be significant ($\chi^2_1 = 455.00$, $p = 5.89 \times 10^{-101}$, BF > 100), indicating correct recognition was greater than false alarms to unstudied buffer items (.69 vs. .16), confirming participants successfully discriminated between studied and unstudied items. Response times for false alarms were also analysed via linear mixed-effects models. Again, random effects for test word in addition to random effects for subjects were found to be necessary ($\chi_1^2 = 186.70$, $p = 1.06 \times 10^{-42}$, BF > 100). Mean response times were not found to differ significantly between studied buffer hits and unstudied buffer item false alarms ($\chi_1^2 = 0.40$, $p = .53$, BF = 0.01).

General discussion

The primary goal of this study was to evaluate the time course of CI priming on different speeded response tasks following study of mediated false memory lists. Several important findings emerged. First, short- and long-term semantic priming effects were found in both the SCT and naming tasks when the CI was presented in the early, middle, and late test positions. Although CI semantic priming was reliable across conditions, a decay pattern emerged in which priming was greatest when CIs were presented in the first test position, but declined across remaining test positions, though the priming effect remained reliable. Significant CI priming effects across test positions in the SCT and naming tasks mapped on to the reliable false recognition patterns that were found across test positions. This correspondence between priming and false recognition is consistent with notion that automatic spreading activation processes contribute to false recognition, even in relatively long-term priming contexts.

Our study follows up on prior work, which has also generally found evidence for long-term CI priming with directly related DRM lists in which response tasks were completed within 1 min following study (Hancock et al., 2003; Meade et al., 2010; Tse & Neely, 2005, 2007). As noted in the Introduction, our use of mediated lists to gauge long-term priming comes with notable benefits. First, the implicit nature of mediated lists means that any activation of CIs must occur through some low-level automatic spreading activation process, which is unlikely to be supported by gist extraction processes. With DRM lists, the strong association between list items and the CI also produces a coherent theme for each list which may also contribute to previously documented long-term priming effects. Mediated lists do not contain a strong list theme, which eliminates this possibility from contention (cf. Brainerd & Reyna, 2002). Second, post-lexical strategies, which may befall speeded response tasks with related study materials, are less likely to occur with mediated lists as these lists are perceived as unrelated and participants have been unable to successfully identify CIs under explicit instructions (Coane et al., 2016; Huff et al., 2012). The mediated false memory paradigm therefore provides an ideal method for gauging the time course of CI activation while minimising shortcomings of speeded response tasks following study of DRM lists.

The shape of the mediated priming effect across test positions found in our study is also consistent with general semantic priming effects reported in the literature. Specifically, the finding that CI priming was robust in the first test position but then declined in later test positions aligns with findings that semantic priming effects are greatest immediately following the prime, but then quickly dissipate. One could argue perhaps that our long-term priming effects are inconsistent with other priming studies, which reported that priming is eliminated after a single intervening item (Dannenbring & Briand, 1982; Masson, 1995). It is important to clarify, however, that in traditional priming experiments, participants are exposed to a single prime followed immediately by a target. In contrast, participants in our experiment (and in Meade et al., 2007; Tse & Neely, 2005, 2007) are exposed to many "primes" (i.e., DRM list items) before responding to the target CI. The presentation of several primes may, therefore, contribute to long-term priming activation, even if priming is more robust when the CI target is presented early in the test list. In contrast to our predictions, however, priming patterns were equivalent between the SCT and naming tasks. Based on previous work that has shown the SCT to be more sensitive to semantic effects in visual word recognition (e.g., Yap et al., 2012), we expected that priming effects, including any long-term priming, would be greater for the SCT over the naming. Our finding of task equivalence, therefore, indicates that long-term priming of mediated CIs under conditions that minimise strategic lexical processes is reliable and generalises over at least two speeded tasks with keypress and vocal responses.

Of course, the low level of association to CIs with mediated lists produced far lower rates of false recognition than is typically found using DRM materials. For instance, Meade et al. (2007) reported a DRM false recognition effect (false alarms to DRM CIs minus false alarms to controls) ranging between 30% and 40% across experiments. In contrast, the corrected false recognition effect of mediated CIs in our experiment was considerably smaller, ranging between 5% and 6% across positions. The lower false recognition rates for mediated CIs are unsurprising given the indirect associations (see Coane et al., 2016; Huff et al., 2012, for direct comparisons between mediated and DRM lists), but the false recognition effect was stable across test positions, which is an important similarity to the patterns reported in studies that used DRM lists above. Thus, despite the overall lower rates of false recognition for mediated CIs, priming effects can still be measured using speeded response tasks.

The long-term mediated priming effects provide further support for automatic spreading activation processes, a critical mechanism in AMT that is used to account for the DRM illusion (Roediger, Balota, and Watson, 2001). As reviewed above the DRM paradigm confounds AMT with thematic gist extraction in FTT (Brainerd & Reyna, 2002). However, the indirect nature of mediated lists prevents extraction of a consistent list theme, leaving a "purer" method to gauge activation processes. In DRM priming studies, in addition to spreading activation, it is possible that priming could reflect a persistent gist trace formed at study. However, the mediated lists used in our study do not contain a consistent theme which eliminates the availability of a consistent gist trace. Given that even when participants are instructed on how mediated list items are related to CIs and asked to explicitly guess CIs, they are rarely able to do so successfully (Coane et al., 2016; Huff et al., 2012), providing further evidence that a consistent gist trace is unavailable.

Given our focus on implicit processes, our experimental procedures were careful to restrict contributions of explicit processes. In addition to the use of tasks that minimise semantic matching, a response deadline of 1,000 ms was also included. This deadline was important as it encouraged participants to remain on task during the experiment to ensure that implicit activation from the study list was still present. These procedures were necessary given that spreading-activation processes are generally short-lived. We do note that our response deadline was shorter than the upper cutoff used in Meade et al.'s (2010) speeded naming task. However, latencies between our experiment and Meade et al.'s were quite similar (position means ranging between 494 and 561 ms in our experiment and 452 and 549 ms in Meade et al.), indicating that response rates were equivalent regardless of whether a response deadline was present or not. Similarly, a response deadline may have also precluded participants from utilising explicit recollection processes when making recognition decisions (e.g., Jacoby, 1991).

In addition, it is worth noting that although we have contextualised our experiments as aligning with other "semantic" priming studies, the mediated lists themselves are not semantically related in the sense that they share a consistent theme/meaning. Instead, they are indirectly associatively related, which by definition means there is not a strong meaning-based gist. In semantic priming paradigms, the term "semantic" typically accounts for both associative and semantic processes, which is sensible given primes and targets are generally both associatively and thematically related (see Hutchison, 2003, for discussion of semantic/associative distinctions). Given the lack of consistent semantic themes across mediated list items, we acknowledge here that mediated CI activation is a result of associative-based rather than semantic-based processes.

Finally, we also acknowledge that while our study was designed to examine long-term semantic priming effects of mediated lists, the delay between study and the CI even in the last test position is relatively short (seven intervening test items). Hancock et al. (2003) and Tse and Neely (2005), who similarly found evidence for long-term CI priming, included a 30-s filler task after prior to the test phase which was a much longer delay than used in our study. We chose not to include a filler delay given mediated activation is indirect and likely smaller than activation from directly associated DRM items. Thus, we reasoned that if CI priming effects would occur, they would be more likely to emerge immediately following study rather than following a filler task. Whether mediated CI priming operates similarly to DRM lists and is detected after a filler is an untested question, but we argue that even a delay of a few intervening items is lengthy, given that participants are also exposed to several semantically unrelated items. Indeed, most priming studies utilise very short delays and do not examine the correspondence between priming and subsequent memory. Our evidence for long-term semantic priming supports the notion that low-level implicit activation is associated with long-term memory processes.

Conclusion

This study was designed to evaluate long-term semantic activation of CIs that originate from mediated false memory lists. Our results showed reliable CI priming across test positions using SCT and speeded naming tasks, consistent with long-term priming effects, and evidence for priming decay with priming greatest in the first test position, which was the most recent test position following study. Long-term priming effects followed false recognition patterns in which false alarms to mediated CIs were reliable and stable across test positions. The use of mediated lists to test for long-term priming is noteworthy as mediated false memories have been argued to occur only through implicit spreading activation processes (Huff et al., 2012) and post-lexical strategies, which are more likely to affect responses when primes and targets are directly related, are less likely to occur. Our study therefore yields stronger evidence for long-term effects of automatic spreading activation that operate when post-lexical strategies are minimised.

Acknowledgements

The authors wish to thank the following research assistants for their support during data collection: Kai Chang, Tamar Cimenian, Sam Gray, Bridget Horwood, Kaitlin McManus, Cole Walsh, and Shuofeng Xu.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding for this study was provided by an Eagle Scholars Program for Undergraduate Research Grant at The University of Southern Mississippi to A.D.M. Additional funding was provided by a James McDonnell Foundation *Understanding Human Cognition* Grant awarded to J.H.C. (#220020426).

ORCID iD

Mark J Huff https://orcid.org/0000-0002-0155-7877

Data accessibility statement

The experiment reported here was not preregistered. All data and stimuli in this study are available to other researchers upon request.

Notes

1. Such strategies refer to situations in which it is posited that participants rely, intentionally or not, on prior encoding episodes to facilitate the lexical decision. For example, on the lexical-decision task a participant might think back to the prime and search for a relationship with the target to bias a word response. Such a strategy would result in facilitation for related pairs, but would not necessarily reflect a forward-acting spread of activation from the prime to the target. Similarly, participants could generate potential targets upon prime presentation, especially if the type of relations between prime and target pairs is predictable (e.g., if many prime–target pairs are antonyms, participants could strategically generate antonyms in a proactive manner). Thus, using methods to reduce possible strategy use will provide a more accurate measure of spreading activation processes.
2. Although the power analysis supports our sample size to detect the predicted small effects, Brysbaert and Stevens (2018) have recommended that a sufficiently powered repeated-measures reaction time experiment includes at least 1,600 observations per condition. With only four observations per condition for each participant, our experiment only approximates this number when collapsed across test position. However, we note that our sample size meets and even exceeds many of the experiments on which our study is based (e.g., Meade et al., 2007, 2010; Tse & Neely, 2005, 2007), providing confidence that the sample size used is sufficient to detect the expected priming effects.

References

Abadie, M., & Camos, V. (2019). False memory at short and long term. *Journal of Experimental Psychology: General*, *148*, 1312–1334. https://doi.org/10.1037/xge0000526

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

Balota, D. A., Black, S. R., & Cheney, M. (1992). Automatic and attentional priming in young and older adults: Reevaluation of the two process model. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 485–502. https://doi.org/10.1037/0096-1523.18.2.485

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 340–357. https://doi.org/10.1037/0096-1523.10.3.340

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

Balota, D. A., & Paul, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 827-845. https://doi.org/10.1037/0278-7393.22.4.827

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . Trieman, R. (2007). The English lexicon project. *Behavioral Research Methods*, *39*, 445–459. https://doi.org/10.3758/BF03193014

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, *49*, 229–240. https://doi.org/10.1080/00221309.1953.9710088

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, 10, 3–47. https://doi.org/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. https://doi.org/10.1110/1467-8721.00192

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. https://doi.org/10.3758/PBR.15.6.1035

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improve word frequency measure for American English. *Behavioral Research Methods, Instruments & Computers*, *41*, 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*, 1–20. https://doi.org/10.5334/joc.10

Cann, D. R., McRae, K., & Katz, A. N. (2011). False recall in the Deese-Roediger-McDermott paradigm: The roles of gist and associative strength. *Quarterly Journal of Experimental Psychology*, *64*, 1515–1542. https://doi.org/10.1080/17470218.2011.560272

Castel, A. D., McCabe, D. P., & Roediger, H. L., III. (2007). Illusions of competence and overestimation of associative memory for identical items: Evidence from judgments of learning. *Psychonomic Bulletin & Review*, *14*, 107–111. https://doi.org/10.3758/BF03194036

Chwilla, D. J., & Kolk, H. H. J. (2002). Three-step priming in lexical decision. *Memory & Cognition*, *30*, 217–225. https://doi.org/10.3758/BF03195282

Chwilla, D. J., Kolk, H. H. J., & Mulder, G. (2000). Mediated priming in the lexical decision task: Evidence from event-related potentials and reaction time. *Journal of Memory and Language*, *42*, 314–341. https://doi.org/10.1006/jmla.1999.2680

Coane, J. H., Huff, M. J., & Hutchison, K. A. (2016). The ironic effect of guessing: Increased false memory for mediated lists in younger and older adults. *Aging, Neuropsychology, and Cognition*, *23*, 282–303. https://doi.org/10.1080/13825585.2015.1088506

Coane, J. H., & McBride, D. M. (2006). The role of test structure in creating false memories. *Memory & Cognition*, *34*, 1026–1036. https://doi.org/10.3758/BF03193249

Coane, J. H., McBride, D. M., Raulerson, B. A., III, & Jordan, J. S. (2007). False memory in a short-term memory task. *Experimental Psychology*, *54*, 62–70. https://doi.org/10.1027/1618-3169.54.1.62

Colbert, J. M., & McBride, D. M. (2007). Comparing decay rates for accurate and false memories in the DRM paradigm. *Memory & Cognition*, *35*, 1600–1609. https://doi.org/10.3758/BF03193494

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, *82*, 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *33*, 497–505. https://doi.org/10.1080/14640748108400805

Dannenbring, G. L., & Briand, K. (1982). Semantic priming and the word repetition effect in a lexical decision task. *Canadian Journal of Experimental Psychology*, *36*, 435–444. https://doi.org/10.1037/h0080650

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146

Gallo, D. A. (2006). *Associative Memory Illusions*. New York: NY: Psychology Press.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, *38*, 833–848. https://doi.org/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. https://doi.org/10.3758/BF03196194

Hancock, T. W., Hicks, J. L., Marsh, R. L., & Ritschel, L. (2003). Measuring the activation level of critical lures in the Deese-Roediger-McDermott paradigm. *The American Journal of Psychology*, *116*, 1–14. https://doi.org/10.2307/1423332

Hill, H., Strube, M., Roesch-Ely, D., & Weisbrod, M. (2002). Automatic vs. controlled processes in semantic priming differentiation by event-related potentials. *International Journal of Psychophysiology*, *44*, 196–218. https://doi.org/10.1016/S0167-8760(01)00202-1

Huff, M. J., & Bodner, G. E. (2013). When does memory monitoring succeed versus fail? Comparing item-specific and relational encoding in the DRM paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1246–1256. https://doi.org/10.1037/a0031338

Huff, M. J., & Bodner, G. E. (2014). All varieties of encoding variability are not created equal: Separating variable processing from variable tasks. *Journal of Memory and Language*, *73*, 43–58. https://doi.org/10.1016/j.jml.2014.02.004

Huff, M. J., & Bodner, G. E. (2019). Item-specific and relational processing both improve recall accuracy in the DRM paradigm. *Quarterly Journal of Experimental Psychology*, *72*, 1493–1506. https://doi.org/10.1177/1747021818801427

Huff, M. J., Coane, J. H., Hutchison, K. A., Grasser, E. B., & Blais, J. E. (2012). Interpolated task effects on direct and mediated false recognition: Effects of initial recall, recognition, and the ironic effect of guessing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1720–1730. https://doi.org/10.1037/a0028476

Huff, M. J., & Hutchison, K. A. (2011). The effects of mediated word lists on false recall and recognition. *Memory & Cognition*, *39*, 941–953. https://doi.org/10.3758/s13421-011-0077-0

Huff, M. J., McNabb, J., & Hutchison, K. A. (2015). List blocking and longer retention intervals reveal an influence of gist processing for lexically ambiguous critical lures. *Memory & Cognition*, *43*, 1193–1207. https://doi.org/10.3758/s13421-015-0533-3

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A micro analytic review. *Psychonomic Bulletin & Review*, *10*, 785–813. https://doi.org/10.3758/BF03196544

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. https://doi.org/10.1016/j.jml.2004.08.003

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

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

Jones, L. L. (2012). Prospective and retrospective processing in associative mediated priming. *Journal of Memory and Language*, *66*, 52–67. https://doi.org/10.1016/j.jml.2011.08.005

Jones, L. L., & Estes, Z. (2012). Lexical priming: Associative, semantic, and thematic influences on word recognition. In J. S. Adelman (Ed.), *Visual word recognition: Meaning and context, individuals and development* (Vol. *2*, pp. 44–72). Psychology Press.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1083–1105. https://doi.org/10.1037/0278-7393.23.5.1083

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. https://doi.org/10.1080/01621459.1995.10476572

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 187–194. https://doi.org/10.1037/0278-7393.31.2.187

Lampinen, J. M., Neuschatz, J. S., & Payne, D. G. (1999). Source attributions and false memories: A test of the demand characteristics account. *Psychonomic Bulletin & Review*, *6*, 130–135. https://doi.org/10.3758/BF03210820

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. https://doi.org/10.1080/09658210244000404

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23. https://doi.org/10.1037/0278-7393.21.1.3

McCabe, D. P., & Smith, A. D. (2002). The effect of warnings on false memories in young and older adults. *Memory & Cognition*, *30*, 1065–1077. https://doi.org/10.3758/BF03194324

McKone, E. (2004). Distinguishing true from false memories via lexical decision as a perceptual implicit test. *Australian Journal of Psychology*, *56*, 42–49. https://doi.org/10.1080/00049530410001688092

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.

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

Meade, M. L., Hutchison, K. A., & Rand, K. M. (2010). Effects of delay and number of related list items on implicit activation for DRM critical items in a speeded naming task. *Journal of Memory and Language*, *62*, 302–310. https://doi.org/10.1016/j.jml.2009.11.009

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

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234. https://doi.org/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. https://doi.org/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. https://doi.org/10.1037/0096-3445.106.3.226

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Lawrence Erlbaum.

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

Psychology Software Tools. (2016). E-Prime 3.0. http://www.pstnet.com

Rabinowitz, J. C., Craik, F. I. M., & Ackerman, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology*, *36*, 325–344. https://doi.org/10.1037/h0080643

Ray, S., & Bly, B. M. (2007). Investigating long-term semantic priming of middle- and low-familiarity category exemplars. *Journal of General Psychology*, *134*, 453–466. https://doi.org/10.3200/GENP.134.4.453-466

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narrative. *Journal of Applied Research in Memory and Cognition*, *5*, 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Robinson, K. J., & Roediger, H. L., III. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*, 231–237. https://doi.org/10.111/j.1467-9280.1997.tb00417.x

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

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814. https://doi.org/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. https://doi.org/10.3758/BF03196177

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (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. https://doi.org/10.3758/BF03194323

Spreen, O., & Benton, A. L. (1977). *Neurosensory center comprehensive examination for aphasia: Manual of directions* (Rev. ed.). Neuropsychology Laboratory, University of Victoria.

Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, *29*, 424–432. https://doi.org/10.3758/BF03196393

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. https://doi.org/10.1016/j.jml.2005.07.001

Tse, C.-S., & Neely, J. H. (2007). Semantic and repetition priming effects of Deese/Roediger-McDermott (DRM) critical items and associates produced by DRM and unrelated study lists. *Memory & Cognition*, *35*, 1047–1066. https://doi.org/10.3758/BF03193477

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

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 1–10. https://doi.org/10.3389/fnhum.2012.00072

Zeelenberg, R., & Pecher, D. (2002). False memories and lexical decision: Even twelve primes do not cause long-term semantic priming. *Acta Psychologica*, *109*, 269–284. https://doi.org/10.1016/S0001-6918(01)00060-9