THE INFLUENCE OF SEMANTIC CONTEXT ON FALSE MEMORIES

by

BRETT HUNTER BALL

(Under the Direction of Adam Goodie)

ABSTRACT

Although false memories have largely been examined with the Deese-Roediger-McDermott (DRM) paradigm, little research has focused on the semantic context in which associates are encoded. Across three experiments, we varied semantic context during a sentential processing task with DRM associates embedded within sentences. More meaningful sentences resulted in greater memory errors (Experiment 1). Furthermore, providing contextual information to discriminate old from new items did not reduce false alarms relative to encoding words in isolation when sentences converged on the meaning of the critical lure (Experiment 2A), and actually increased memory errors (Experiment 3). We provide evidence for lure activation at encoding (Experiment 2B) and suggest an important role of reactivation during retrieval. These results suggest that semantic context that allows for meaningful organization of items within-lists (internal convergence) and the ability for individual items to form associative connections to the critical lure (external convergence) increases false memories.

INDEX WORDS: False memory, Recognition memory, DRM, Spreading activation, Reactivation

THE INFLUENCE OF SEMANTIC CONTEXT ON FALSE MEMORIES

by

BRETT HUNTER BALL

BS, University of Georgia, 2009

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2011

THE INFLUENCE OF SEMANTIC CONTEXT ON FALSE MEMORIES

by

BRETT HUNTER BALL

Major Professor: Adam Goodie

Committee: Jason Hicks
Nash Unsworth

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2011

DEDICATION

In honor of Richard L. Marsh, for without whom this achievement would not have been possible.

ACKNOWLEDGEMENTS

I thank my Committee, as well as Mike DeWitt and Justin Knight, for their comments on previous versions of this manuscript.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 INTRODUCTION 1

The Present Study 6

2 EXPERIMENT 1 9

Methods 9

Results 11

Discussion 13

3 EXPERIMENT 2A 15

Methods 16

Results 17

Discussion 19

4 EXPERIMENT 2B 21

Methods 22

Results 24

Discussion 27

5 EXPERIMENT 3 29

Methods....................30

Results....................31

Discussion....................33

6 GENERAL DISCUSSION 35

Relational Processing....................38

Activation/Monitoring Theory....................40

Alternative Theories....................43

Conclusion 44

REFERENCES 46

LIST OF TABLES

Page

Table 2.1: Recognition Hit Rates and False Alarm Rates (Standard Errors) by Encoding Block for Experiment 1 12

Table 3.1: Recognition Hit Rates and False Alarm Rates (Standard Errors) by Condition for Experiment 2A 18

Table 4.1: Response Latencies (ms) for Primed and Unprimed Critical Lures (Standard Errors) by Condition for Experiment 2B 25

Table 5.1: Recognition Hit Rates and False Alarm Rates (Standard Errors) by Condition for Experiment 3 32

LIST OF FIGURES

Page

Figure 4.1: Priming effects for critical lures by condition 26

Figure 4.2: Priming effects for associates by condition 27

CHAPTER 1

Introduction

Examination of the factors involved in the creation of false memories provides a fruitful method of investigating the underlying mechanisms involved in the organization of human memory. Memory errors have been largely studied using a variety of methods and have been found to occur when new words or sentences are similar in meaning to studied items (e.g., Bransford & Franks, 1971), for new items that are visually or phonologically similar to old items (e.g., Jones & Jacoby, 2001), during free recall tests (Unsworth & Brewer, 2010), and during eyewitness testimony (e.g. Loftus, 1971). A method that has been widely used to investigate semantic false memories is the Deese-Roediger-McDermott (DRM) paradigm whereby after studying a list of related words such as *bed, rest, tired,* and *dream*, people often erroneously claim that a non-presented critical lure (*sleep*) was originally studied (Deese, 1959; Roediger & McDermott, 1995). The current study was designed to examine semantic false memories in the DRM recognition memory paradigm by varying semantic contexts in sentential processing tasks during encoding.

Central to the underlying mechanisms involved in false memories using the DRM paradigm is the organization of list items. Each item within a list is associated (by frequency of co-occurrence) with a common theme, which can include orthographic, categorical, or conceptual similarities between the items and theme. The typical DRM paradigm presents list items in a blocked fashion according to theme and has participants try to remember these items for a later memory test. Relating items in terms of categorical or conceptual relationships benefits veridical memory performance but may also increase susceptibility of false memory errors (McCabe, Presmanes, Robertson, & Smith, 2004). The degree to which relationships between items are formed may allow extraction of the overall theme of the associated list or cause the critical lure to become activated and available in memory, depending on the subjective organization imposed by the participant during encoding (Gallo, 2006).

The cognitive processes underlying the DRM false memory paradigm are not fully understood. One theory that has been proposed to explain the DRM illusion is fuzzy-trace theory (Brainerd & Reyna, 2002; Reyna & Brainerd, 1995), which suggests that encoding of list items results in two types of memory traces: verbatim and gist. Verbatim traces include specific contextual details from processing surface forms of experienced items whereas gist traces reflect extracted commonalities among experiences whereby participants mentally construct a gist representation of common features of the conceptual form of items. Because no verbatim traces are present for critical lures, this theory proposes that false alarms to critical lures reflect the match of the critical lure and the gist representation from studied associates based on its familiarity (or in some instances, "phantom recollection"; Brainerd, Wright, Reyna, & Mojardin, 2001). An alternative theory used to explain the creation of false memories is activation/monitoring theory (Roediger, Watson, McDermott, & Gallo, 2001). According to this theory, false alarms to critical lures are due to processes that internally activate the conceptual representation of the critical lure (Gallo, 2010), whereas false alarms are decreased due to decision processes that monitor for the origin of the memory signal. Although considerable evidence has provided support for (and against) both theories, the present study was designed to assess the predictions of the activation/monitoring theory.

Activation/monitoring theory posits that activation of critical lures occurs during processing of list items via spreading activation of conceptual representations within a semantic network (Anderson, 1983; Collins & Loftus, 1975; Gallo, 2010). During encoding, a summation of multiple implicit associative responses produced by the studied associates may internally activate the conceptual representation of the critical lure thus making it available in memory (Hancock, Hicks, Marsh, & Ritschel, 2003; Underwood, 1965). During retrieval, test probes may serve to reactivate the associative network that subsequently makes the critical lure susceptible to false remembering due to a high degree of overlap between the lure and its activated representation within the associative network. (Kimball, Muntean, & Smith, 2010; Meade, Watson, Balota, & Roediger, 2007). However, monitoring processes at retrieval can serve to reduce false memories (Unsworth & Brewer, 2010). The source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993) has been applied to explain how qualitative characteristics of a memory trace are evaluated and retrieved. For example, experienced events (i.e. studied items) should contain more perceptual details relative to internally generated events (i.e. critical lures) which should include more details associated with cognitive operations. The degree to which features of the critical lure resembles actual studied items increases the probability that false recall or recognition will occur due to a reality-monitoring error in which participants mistake internally generated items as actually being perceived (e.g., Hicks & Marsh, 1999, 2001). Thus, within the activation/monitoring theory, critical lures become highly activated (or reactivated) due to spreading activation during encoding or retrieval and false memories occur when monitoring processes fail.

Manipulations that influence semantic processing directly have been shown to increase both veridical and false memory. Toglia, Neuschatz, and Goodwin (1999) found that in addition to greater true recall, semantic processing led to greater false recall of critical lures and that blocking lists of associates increased false memories. Other studies examining varying levels-of-processing (Craik & Lockhart, 1972) have shown that deeper processing increases false memories (Rhodes & Anastasi, 2000; Thaper & McDermott, 2001), as do instructions that encourage relational processing (McCabe et al., 2004). It is suggested that semantic processing results in greater activation of related items within an associative network due to the strengthening of semantic relationships among items and that blocking list presentation highlights the common semantic features within the studied list making it more likely that the critical lure is activated. However, other encoding manipulations have been shown to reduce the occurrence of false memories. When distinctive pictorial information is encoded with the auditory presentation of studied DRM items, reductions in false alarms to critical lures occur relative to encoding of aurally presented words along with visually presented words (Israel & Schacter, 1997). Similarly, distinctive processing at encoding has been shown to reduce false alarms to critical lures when list items are presented in unique fonts (Arndt & Reeder, 2003), as well as instructions that focus attention on item-specific features of each list item (McCabe et al., 2004). Goodwin, Meissner, and Ericsson (2001) found that manipulating the contextual organization of study lists by intermixing unrelated filler items with semantic associates resulted in a reduction of false memories because making semantic connections to the critical lure was more difficult. Thus, according to activation/monitoring theory strengthening semantic relationships between studied items through relational processing may increase the probability that the critical lure is activated, whereas decreasing list organization or increasing item-specific processing may decrease spreading activation of the studied items to the related lure resulting in a weaker memory signal.

To date, only a handful of DRM studies have provided contextual information during encoding by placing associates in the context of sentences or text. Dewhurst, Pursglove, and Lewis (2007) found that 5-year-olds were more likely to falsely recognize critical lures after reading stories with DRM associates placed within the sentences of the text as compared to the standard list encoding (although there were no differences for 8 and 11-year-olds). They argued that the story context made it easier for younger children to identify its overall theme, whereas older children may have a greater ability to identify thematic associations during list processing and therefore had equally high false alarms in both encoding conditions. Plancer, Nicolas, and Piolino (2008) examined the influence of suggestibility in younger adults after participants read a text with DRM associates placed within the sentences. If participants did not recall the critical lure after reading the text, the experimenter suggested that the critical lure was present and examined false alarms to critical lures on a subsequent recognition test. False recognition was greater with strong relative to no suggestion, and similar results were found when encoding only list items. Although a cross-experimental comparison was not conducted, results suggested that text processing did not increase susceptibility to false recognition relative to list processing. Comparing younger and older adults, Thomas and Sommers (2005) found that younger adults were able to use item-specific information to reduce false memories when associates were presented in the context of sentences relative to words encoded in isolation, whereas older adults only showed a similar reduction when the sentence context did not elicit the meaning of the critical lure.

Although the methodologies and age ranges between these studies differ markedly, no clear pattern of results is evident. The results from Plancer et al. (2008) and Dewhurst et al. (at least with older children; 2007) suggest that placing DRM associates in the context of sentences does not increase susceptibility of false memories, whereas the results from Thomas and Sommers (2005) suggest that sentence processing actually decreases false memories, relative to list processing. One possibility is that in all three experiments, standard list processing may have encouraged relational processing. However, in the former two studies reading texts (rather than individual sentences) may have encouraged comparable relational processing to the standard list encoding, whereas in the latter study associates were embedded within individual sentences that may have increased item-specific processing. Differences in false alarm rates across experiments could be due to differences in the type of processing at encoding whereby relational encoding may increase semantic relationships between items resulting in greater lure activation and item-specific processing may disrupt this activation.

The Present Study

The primary goal of the present study was to examine how varying the semantic context of sentences influences false memories when DRM associates are placed in the context of sentences. Although semantic processing has been shown to increase memory errors in the DRM paradigm (e.g. Toglia et al., 1999; Goodwin et al., 2001), providing contextual information that differentiates items at encoding can reduce false memories (e.g. Arndt & Reeder, 2003; Thomas & Sommers, 2005). Thus, sentential processing could potentially strengthen semantic relationships among items and increase lure activation or be used to distinguish items from one another. Furthermore, we were interested in determining whether differences in activation could account for differences in memory performance by using an implicit memory test to examine priming effects for the critical lure from previously studied DRM lists.

The majority of the DRM research requires participants to intentionally memorize list items for a later memory test, whereas only a few studies have employed a surprise memory test (e.g. Dodd & MacLeod, 2004; Rhodes & Anastasi, 2000; Tussing & Greene, 1997). Presumably, intentional learning instructions may encourage participants to identify semantic commonalities among items thus making it more likely to identify or activate the critical lure during encoding. In fact, previous research has found that participants sometimes produce the critical lure during rehearsal (Goodwin et al., 2001; Seamon, Lee, Toner, Wheeler, Goodkind, & Birch, 2002). Although research suggests intentional learning is not necessary for the creation of false memories (e.g. Dodd & MacLeod, 2004), we sought to extend previous findings using an incidental learning paradigm during a sentential processing task while also empirically measuring critical lure activation levels using an implicit memory test.

For each experiment, we placed associates in the context of sentences blocked by theme such that the last word in each sentence was an associate of a non-presented critical lure. In Experiment 1, the sentence structure allowed for meaningful processing of list items for half the blocks whereas the other half did not. The sentence context in Experiment 2A converged on the meaning of the critical lure in one condition and diverged from the meaning in another condition, and false alarm rates were compared to a condition in which words were encoded in isolation. Experiment 2B assessed lure activation during an implicit memory test by examining priming effects during a lexical decision task. Experiment 3 used stimuli other than those typically used in the DRM paradigm that converged on two different meanings of a homographic lure (Hutchinson & Bolota, 2005) to compare recognition of items studied in sentences to words studied in isolation. We hypothesized that the degree to which sentential processing allows for meaningful organization of items within a list and semantic connections to the critical lure, increases in false alarm rates should be seen relative to sentence contexts that disrupts these processes. Furthermore, we predicted that when sentence processing elicits the meaning of the critical lure, false alarm rates would increase relative to processing words in isolation. To foreshadow our results, such sentence processing did not uniformly increase false alarms relative to word processing across experiments.

CHAPTER 2

Experiment 1

In Experiment 1 we investigated the influence of semantic processing on false memories by placing DRM associates in the context of sentences where for half of the sentence blocks the structure allowed for meaningful processing of sentences (e.g. "John visited the hospital.") whereas the other half did not (e.g. "Sara drank the desk."). For clarity, we will refer to the former type as "meaningful" and the latter as “meaningless”. We predicted that false alarms would be greater for meaningful blocks due to greater semantic and/or relational processing of items within a list that may be disrupted in the meaningless blocks. Furthermore, if participants are engaging in item-specific processing during meaningless sentences due to disrupted relational processing (Hege & Dodson, 2004), veridical recognition for meaningless blocks may be greater than meaningful blocks.

Methods

Participants. A total of 30 undergraduate students from the University of Georgia volunteered in exchange for partial credit toward a course research requirement. Each participant was individually tested in sessions that lasted approximately 15 minutes.

Materials. The experimental materials consisted of 16 themed lists taken from the Roediger, Watson, McDermott, and Gallo (2001) norms. We selected eight semantic associates from each list to create sentences for each non-presented critical lure. Eight meaningful and eight meaningless sentences were created for each theme (256 total sentences) for counterbalancing purposes. Other than the verb used, the sentence structure was consistent for both meaningful and meaningless sentences. In the meaningful sentences, the verb used allowed for meaningful processing of the sentence (e.g. “John visited the hospital”) whereas in the meaningless sentences it did not (e.g. “John dealt the hospital). Note that in both sentence types the last word “hospital” is associated to the non-presented critical lure “doctor”.

Design and procedure. The study session was blocked within-subjects, such that a block of eight meaningful sentences from six themed lists alternated with a block of eight meaningless sentences from six others. The sentences within each block were randomly presented, and for half of the participants a meaningful block was presented first while a meaningless block was presented first of the other half. These alternating blocks persisted until 12 blocks (six meaningful and six meaningless) were presented. Items from the other four DRM sentence lists served as new items during the test phase. The counterbalancing scheme served to ensure each list was presented an equal number of times as both meaningful and meaningless sentences across participants, and also served as a pool for new lures during the test phase equally. So for one third of the participants, the “doctor” list was presented in meaningful blocks and in meaningless blocks for another third. Furthermore, for one third of the participants, the “doctor” list was not studied and served as new items during the test phase. This process occurred for each of the 16 lists.

The test phase consisted of 48 old items and 48 new items. Of the old items, 4 were taken from each of the 12 presented lists. The new items consisted of 12 non-presented critical lures (e.g. “doctor”) from each sentence list of the study phase. As described in the counterbalancing scheme, 4 sentence lists were not presented during study. Of these 4 lists, the critical theme word and 4 associates of each (20 items) were presented as new items. In addition, 16 new items were taken from the Roediger et al. (2001) norms that were unrelated to the other items. The 96 items during the test phase were randomly presented.

For each phase of the experiment, participants read the instructions from the computer monitor which the experimenter also reiterated in her own words. The instructions for the study phase indicated that we were interested in seeing how people rated sentences for meaning. Participants were told that they would be presented with a series of sentences and to rate each sentence for subjective meaning on a scale from 1 to 7 (1 being absolutely meaningless, 7 being absolutely meaningful). The presentation rate was self-paced with a 5-second break between each block. Upon conclusion of the study phase, a 2-minute distracter phase consisting of a series of mazes to be solved was administered. Following this, instructions for the surprise recognition test were given. Participants were told they were going to be shown a series of items. Upon presentation, they were to think back to the sentences rated earlier and if they remembered seeing the presented word in one of the sentences they were to press the “yes” key. If the item was new, they were to press the key labeled "no" to indicate that they did not see the word during the previous rating task.

Results

Sentence rating task. Unless otherwise specified, all statistical tests are significant at the conventional 5% probability of a Type I error. To examine whether our encoding manipulation resulted in differences in perceived meaning between sentence types, a simple comparison was conducted for mean rating scores (1-7) for meaningful versus meaningless blocks. Meaningful blocks (M = 5.98) received significantly higher ratings than meaningless blocks (M = 1.98), $t(29) = 26.45$, $p < .05$, suggesting the encoding task was successful in producing differences in perceived meaning.

Recognition. Table 2.1 displays hits, false alarms to critical lures, and different categories of new items. We conducted a 2 (block: meaningful vs. meaningless) x 2 (item type: studied vs. critical lure) Analysis of Variance (ANOVA) with repeated measures on both factors. There was no main effect of block although it approached conventional levels of significance, $F(1, 29) = 3.22$, $p = .08$, $\eta_p^2 = .10$. A main effect of item type was found, $F(1, 29) = 15.28$, $p < .01$, $\eta_p^2 = .35$, with greater recognition for studied items (M = .74) than critical lures (M = .63). In addition, there was a significant interaction of block and item type, $F(1, 29) = 4.44$, $p < .05$, $\eta_p^2 = .133$. False alarms to critical lures were greater in meaningful blocks (M = .68) than meaningless blocks (M = .57), $t(29) = 2.19$, $p < .05$. However, there were no differences the hit rates for meaningful (M = .74) and meaningless blocks (M = .73), $t(29) = .351$, $p = .728$. It should also be noted that false alarms to critical lures in both the meaningful and meaningless blocks were greater than false alarms to new themes (M = .23) not related to previously studied lists, $t(29) = 7.82$, and 6.17, $p < .001$, respectively. However, there were no differences in false alarm rates to the different categories of new items, $F(1, 28) = 1.54$, $p = .231$, $\eta_p^2 = .099$.

Table 2.1
Recognition Hit Rates and False Alarm Rates (Standard Errors) by Encoding Block for Experiment 1

	Hits	False Alarms			
Block	Studied	CL	New Theme	New Exemplar	New Unrelated
Meaningful	.74 (.02)	.68 (.04)			
Meaningless	.73 (.03)	.57 (.04)			
Overall			.23 (.04)	.23 (.02)	.19 (.03)

Discussion

The primary goal of Experiment 1 was to determine how varying semantic context influences false memories in the DRM paradigm. Processing of meaningful sentences increased false memories relative to meaningless sentences, whereas there were no differences between the two types of sentences in veridical memory. One possibility is that meaningful sentences produced relational processing that was disrupted in meaningless sentences, thus causing participants to focus more on item-specific processing during the latter. Item-specific processing has been shown to reduce false memories relative to relational processing, whereas no differences occur between the two types of processing for veridical memory (McCabe et al., 2004). It could be argued that veridical memory should be better when using item-specific processing; however, relational processing can serve as an effective means of discriminating between old and new items by responding "old" to items that are consistent with the gist of the studied items at the cost of increased false alarm rates. Consistent with this idea, critical lures had a higher false alarm rate when presented in the context of meaningful sentences presumably because they were consistent with the gist of the other studied items. However, meaningless sentences induced more item-specific processing allowing for more effective monitoring and reductions in false memories.

Alternatively, processing of meaningless sentences may have reduced the amount of semantic information extracted by the participants thereby causing a reduction in false alarms. This idea is supported by differences in ratings scores between the two types of sentences. With decreased semantic processing in the meaningless blocks, the representation of the critical lure may not have been as strongly activated as with the meaningful sentences. Therefore, source discrimination was more effective in the meaningless blocks. However, the current experimental design cannot directly address this alternative hypothesis. We also cannot assertively conclude whether relational processing was the sole contributor to the increased false alarm rate, or whether it was a combination of both semantic extraction and relational processing that produced these effects. Therefore, Experiment 2A was designed to examine whether relational processing will increase false alarms to critical lures while holding semantic processing constant.

CHAPTER 3

Experiment 2A

Thomas and Sommers (2005) placed DRM associates in the context of sentences that converged on or diverged from the meaning of a non-presented critical lure and found that in younger adults false alarms were significantly reduced in both sentence types relative to encoding words in isolation, whereas a similar reduction in older adults was only found with divergent sentences. Presumably, older adults were only able to use item-specific contextual information to reduce false alarms when shared cues relating list items together were made less accessible in the divergent condition, whereas younger adults were able to access this information regardless of whether relational cues were available. Therefore, we designed a similar experiment as Thomas and Sommers using a between-subjects design with convergent, divergent, and word-only conditions. Because each sentence or word is (presumably) equally as meaningful across blocks between conditions but relational processing is differentially affected, differences in false alarm rates may be interpreted more precisely. Because convergent sentences elicit the meaning of the critical lures and are all related to one another within each list, we hypothesized that false alarms to critical lures would be greater in the convergent relative to the divergent condition in which sentences do not elicit the meaning of the critical lures and are unrelated to each other within each list. Divergent sentences should reduce relational processing of sentences within a list during encoding and therefore minimize the production of shared cues (Thomas & Sommers, 2005). If subjective organization based on the meaning of the stimuli is responsible for the false alarm rates, false alarms in the convergent condition should be greater than (or at least equal to) the word-only condition due to relational organization of items within a list. However, if participants are able to use item-specific contextual information, higher false alarm rates should be seen in the word-only condition relative to both sentence conditions.

Methods

Participants. Undergraduate students from the University of Georgia volunteered in exchange for partial credit toward a course research requirement. Each participant was individually tested in sessions that lasted approximately 20 minutes. Ninety new participants were randomly assigned to one of the three between-subjects conditions (30 in each). One participant from the divergent condition was removed due to false alarm rates exceeding hit rates, resulting in only 29 participants in this condition. However, the removal of this participant did not significantly affect any critical analyses.

Materials. A total of 12 themed lists with 8 sentences in each were created with identical non-presented critical lures for each condition. The materials for the convergent and divergent sentences were borrowed from Thomas and Sommers (2005). However, we slightly altered the sentences by trying to equate sentence length and eliminating proper nouns. Convergent sentences elicited the meaning of the semantic associates and converged on the meaning of the non-presented critical lure (e.g. "After work he laid down in bed"). Divergent sentences elicited a particular meaning of the associate at the end of each sentence, but did not converge on the meaning or gist of the non-presented critical lure (e.g. "She walked along the river bed"). Note that the last word of both types of sentences is an associate of the DRM theme word "sleep." In the word condition, the same associates were used as in the convergent and divergent conditions (e.g. "bed") but were presented in isolation (i.e. no sentences).

Design and procedure. The procedure used in Experiment 2A was similar to Experiment 1, except that the 12 blocks were randomly presented during the study phase with each sentence (or word) within a block presented randomly. Instructions for the study and test phase in each condition were identical to those given in Experiment 1. After making meaningfulness ratings on all 12 blocks of sentences or words, participants engaged in a 2-minute distractor phase and then were given instructions for the test phase.

The test phase consisted of 48 old items and 48 new items randomly presented. In each condition, 4 old items were taken from each of the studied lists. The new items in all conditions consisted of the 12 non-presented critical themed items, as well as 4 associates from 4 themes that were never studied along with the critical lure from each. There were also 16 unrelated new items taken from other DRM lists.

Results

Sentence rating task. During the encoding task, there were no differences in ratings for meaning across conditions, $F(2, 86) = .29$, p = .752, suggesting that the stimuli in one condition were not perceived as any more "meaningful" than another condition. Furthermore, there were no differences in ratings between any conditions in subsequent Experiments and will therefore not be discussed further.

Recognition. Table 3.1 displays hits, false alarms to critical lures, and false alarms to new unrelated items. Due to differences in false responding to unrelated new items across conditions ($F(2, 86) = 38.34$, $p < .001$, $\eta_p^2 = .471$), we employed a correction for veridical recognition by subtracting the false alarm rates to unrelated items from the hit rates (see Kensinger & Schacter 1999; Thomas & Sommers, 2005). For false recognition, we subtracted the false alarm rate to unrelated items from the false alarm rate to critical lures (see Table 3.1). We conducted a 2 (item type: studied vs. critical lure) x 3 (condition: convergent vs. divergent vs. word-only) mixed ANOVA. The analysis of corrected hit and false recognition scores revealed a main effect of item type, $F(1, 86) = 91.94$, $p < .001$, $\eta_p^2 = .517$, whereby critical lures (M = .41) were recognized less than studied items (M = .61). A main effect of condition was also found, $F(2, 86) = 118.68$, $p < .001$, $\eta_p^2 = .734$. Participants in the divergent condition (M = .26) recognized fewer items than both the convergent (M = .55) and word-only (M = .72) conditions. These main effects were qualified by a significant interaction, $F(2, 86) = 19.74$, p < .001, $\eta_p^2 = .315$.

Table 3.1
Recognition Hit Rates and False Alarm Rates (Standard Errors) by Condition for Experiment 2A

	Hits		False Alarms		
Condition	Studied	Corrected	CL	Corrected	Unrelated
Convergent	.76 (.02)	.56 (.02)	.73 (.03)	.53 (.03)	.20 (.02)
Divergent	.69 (.02)	.37 (.02)	.47 (.04)	.15 (.03)	.32 (.02)
Word-only	.97 (.01)	.89 (.02)	.63 (.04)	.55 (.03)	.08 (.01)

Separate ANOVAs were conducted for corrected hit and false alarm rates across conditions. There was a significant difference in studied items recognized across conditions, $F(2, 86) = 136.1$, p < .001, $\eta_p^2 = .76$. Planned comparisons revealed that participants in the word-only condition (M = .88) recognized more studied items relative to the convergent condition (M = .56; $t(58) = 12.57$, p < .001), whereas fewer items were recognized in the divergent condition (M = .37) than the convergent condition, $t(57) = 5.56$, p < .001. There was also a significant difference in false alarms to critical lures between conditions, $F(2,86) = 50.41$, p < .001, $\eta_p^2 = .540$. Participants in the divergent condition (M = .15) falsely recognized significantly fewer critical lures than the convergent condition (M = .53; $t(57) = 8.88$, $p < .001$). However, there were no differences between the word-only (M = .55) and convergent conditions, $t(58) = .45$, $p = .66$.

Discussion

The purpose of Experiment 2A was to examine whether differences in false recognition would arise between sentence and word-only encoding conditions. Participants in the word-only condition were much more likely to recognize studied words than either sentence condition, and participants in the convergent condition had better veridical recognition than the divergent condition. Differences in veridical recognition between the word-only and convergent conditions should not be surprising due to the increasing demands of processing and storage (and subsequent remembering) of sentences relative to words encoded in isolation (Thomas & Sommers, 2005). Of critical interest was false recognition of critical lures, which was much lower in the divergent condition relative to the convergent and word-only conditions, which did not differ from each other. One possible explanation of these findings is that participants in the divergent condition were able to use item-specific information in order to reduce false recognition of critical lures. However, veridical recognition was also the lowest in this condition, which suggests that these participants had greater difficulty discriminating old and new items. If participants were engaging in item-specific processing, we might expect greater veridical recognition in this condition (Arndt & Reeder, 2003; McCabe et al., 2004). An alternative hypothesis is that participants in both the convergent and word-only conditions used relational encoding that aided veridical memory at the cost of increased false memories. However, because sentences in the divergent condition were dissimilar from one another within a list and not related to the critical lure, participants were unable to relate sentences together and use shared cues to recognize studied items or falsely recognize critical lures.

Although we did not find evidence as did Thomas and Sommers (2005) that younger adults were able to use item-specific information to reduce false recognition in both sentence conditions as compared to the word-only condition, there were several methodological differences that will be elaborated upon in the General Discussion. Our results support the idea that participants are able to identify thematic associations when lists are organized in such a way to increase relational processing (Dewhurst et al., 2007), with this processing being disrupted when semantic associations within-lists are not available. One residual issue from Experiment 1 is whether differences in lure activation can account for differential false alarms rates across conditions. A disruption of relational processing may result in a lack of activation of the critical lure in the divergent condition. Therefore, in Experiment 2B we sought assess lure activation for each condition.

CHAPTER 4

Experiment 2B

The results from Experiment 2A suggest that presenting DRM associates in sentences that converge on the meaning of the non-presented critical lure increase false alarms to these items relative to sentences that diverge from the meaning. According to the activation/monitoring framework, the differences in false alarms to critical lures could be due to greater activation of these lures during encoding for the convergent and word-only conditions, relative to the divergent condition, thus causing more source-monitoring errors at retrieval. Experiment 2B was designed to examine whether differences in activation levels could be detected by using an implicit memory test to assess priming effects. Explicit memory involves conscious retrieval strategies, whereas implicit memory is manifest through priming effects whereby repeated exposure to perceptual features or access to semantic representations of a stimulus facilitates task performance (e.g., faster response times) without specific awareness of the participant (Tse & Neely, 2005). By using an implicit memory test we can measure activation levels of critical lures when intentional retrieval strategies (i.e., monitoring) are minimized.

Deciding whether a string of letters forms a valid English word in a lexical decision task (LDT) does not require retrieval of specific information from the previous study episode in order to arrive at a correct decision so monitoring strategies to determine the source of stimulus activation is not necessary. By measuring the reaction times to critical lures and associates during a LDT, we can get a relatively pure measure of the activation that is produced by studying a list of DRM associates prior to performing the LDT (Tse & Neely, 2005). That is, items that are highly activated due to prior exposure of a DRM study list may result in a faster response latencies relative to words that have not been previously primed. We can measure priming effects for associates that were actually studied on the DRM list as well as for critical lures that were not studied yet semantically related to the studied DRM items.

A study by Hancock et al. (2003) had participants encode a list of items and perform a LDT after each study list that consisted of 10 words and 10 nonwords. The words consisted of two studied associates, a critical lure, a non-studied control-matched word identical in word frequency, letters, and length to the critical lure, and unrelated new words. Using control-matched words as a baseline measure of reaction times, Hancock et al. found significant priming effects for critical lures that were also faster than responses to studied associates, suggesting that the critical lures were highly activated from studying the DRM list associates. The fact that responses were faster to critical lures than studied associates represents a form of "superadditive priming" whereby the sum of multiple implicit activations of the critical lures produced by the studied associates was greater than the sum of the direct activation of the repeated associate and its weaker implicit activations from other associates (Tse & Neely, 2005). Thus, the conceptual representation of the critical lure was highly activated and available in memory resulting in faster response times. Therefore, we used a similar procedure for the current experiment to examine whether differences in activation could explain the differences in false alarm rates between conditions.

Methods

Participants. Undergraduate students from the University of Georgia volunteered in exchange for partial credit toward a course research requirement. Each participant was individually tested in sessions that lasted approximately 45 minutes. 99 new participants were randomly assigned to one of the between-subjects encoding conditions of either convergent, divergent, or word-only, with 33 participants in each condition.

Materials. The experiment consisted of 12 study-test cycles (sentence rating task, distractor, and LDT). The 12 study lists were identical to those used in the convergent, divergent, and word-only conditions of Experiment 2A. To create primed and unprimed lists, we matched each study list with a non-studied list taken from the Roediger et al. (2001) norms. For example, suppose that participants studied the "sleep" list but not the "chair" list. During the LDT, the "sleep" list would be considered the *primed list*, which was composed of one associate (e.g. "tired"), a critical lure (e.g. "sleep"), a control-matched word ("block"), 7 unrelated words, and 10 pseudohomophonic nonwords (see Tse & Neely, 2005). The *unprimed list* was composed in the same manner, except using "chair" as the critical lure and one associate from its list (e.g. "couch") during the LDT. The primed (and unprimed) associate was always presented after the critical lure to avoid within-list priming of the lure during test. For each 20-item LDT list, the critical lure was presented in the 6th position, the associate was presented in the 12th position, and the control item was presented in the 18th position. The unrelated words and nonwords were randomly assigned to the remaining positions. Once the order of items was created for a block, each participant received the same order. The 20 words that comprised the "sleep" list were first presented followed by the 20 words that comprised the "chair" list resulting in a 40 trial LDT after each study block. The presentation orders for primed and unprimed blocks were counterbalanced to determine whether priming effects decay over time during the LDT. For half the study-test trials, the primed block was presented first whereas for the other half the unprimed block was presented first (in alternating order). The 40 total items for each test block were presented continuously, and we were able to directly measure activation by comparing differences in reaction times to primed and unprimed critical lures and associates.

Design and procedure. For each study-test cycle, participants performed the sentence (or word) rating task from the previous Experiments, a 30-second filler task, and a 40 trial LDT task. The instructions for the sentence (word) rating task were identical to those given in Experiment 2A. For the filler task, participants were told to write down a series of three digit numbers from greatest to least on a piece of paper located beside the keyboard. For the LDT, participants were instructed to decide as quickly and accurately as they could whether a string of letters formed a valid English word by pressing the corresponding key labeled "word" or "nonword". After participants were given the instructions for each task, they performed a brief practice session to become familiar with the procedure. After completing the 12 study-test cycles, participants were dismissed.

Results

Priming effects. Table 4.1 displays response latencies for primed and unprimed critical lures and associates, as well as words. Only correct response latencies within 2.5 standard deviations (SDs) of a given participant's mean were included in the analysis. After the trimming procedure, to avoid biasing results with extreme scores, participants with response latencies to word trials that were 3 SDs from the group mean were excluded from the analysis. This resulted in the removal of 1 participant in the convergent condition and 2 in the word-only condition. We conducted a 2 (priming: primed vs. unprimed) x 2 (item type: critical lure vs. associate) x 2 (block: first vs. second) x 3 (condition: convergent vs. divergent vs. word-only) mixed ANOVA with repeated measures on each factor.

Table 4.1

Response Latencies (ms) for Primed and Unprimed Critical Lures (Standard Errors) by Condition for Experiment 2B

	Primed		Unprimed		
Condition	CL	Associate	CL	Associate	Word
Convergent	589 (15)	598 (13)	615 (16)	618 (15)	621 (13)
Divergent	564 (15)	598 (15)	608 (19)	605 (17)	603 (15)
Word-only	592 (12)	599 (13)	611 (10)	635 (17)	632 (14)

There was no main effect of condition, $F(2,93) = .37$, $p = .689$, $\eta_p^2 = .008$. There was a main effect of priming, $F(1,93) = 33.78$, $p < .001$, $\eta_p^2 = .266$, indicating that reaction times (RTs) for primed items (M = 590 ms) were significantly faster than unprimed items (M = 614 ms). However, this priming effect did not interact with condition, $F(2,93) = .10$, $p = .901$, $\eta_p^2 = .002$. There was a main effect of item type, $F(1,93) = 7.98$, $p < .01$, $\eta_p^2 = .079$, with faster RTs for critical lures (M = 596 ms) than associates (M = 608 ms), but no interaction of item type with condition, $F(2,93) = .64$, $p = .531$, $\eta_p^2 = .014$. Although the main effect was significant at the group level, only in the divergent condition were latencies faster for primed critical lures than primed associates, $t(32) = 4.35$, $p < .001$. There was a significant three-way interaction of priming, item type, and condition, $F(2,93) = 4.43$, $p < .05$, $\eta_p^2 = .087$. This interaction reflects that although RTs for primed critical lures were significantly faster than unprimed lures in all conditions (see Figure 4.1), only in the word-only condition were RTs for primed associates significantly faster than unprimed associates, $t(30) = 3.02$, $p < .01$ (although this effect was marginal in the predicted direction for the convergent condition, $t(31) = 1.91$, $p = .065$; see Figure 4.2). Of primary importance for the current experiment, however, is that we found significant priming effects for primed versus unprimed critical lures but no differences between conditions.

There was also a main effect of block, $F(1,93) = 33.012$, $p < .001$, $\eta_p^2 = .262$, with faster RTs for items in the first block (M = 590 ms) than the second block (M = 615 ms), but no interaction with block and condition, $F(2,93) = .221$, $p = .802$, $\eta_p^2 = .005$. However, there was an interaction of priming and block, $F(1,93) = 8.965$, $p < .01$, $\eta_p^2 = .088$. This interaction reflects that the primed versus unprimed difference (M = 12 ms) was smaller in the first block than the second block (M = 36 ms), but this is likely due faster decay of unprimed items than primed items across blocks. No other interactions were significant. Overall, these results suggest that priming effects decay over time, but this decay does not differ between conditions.

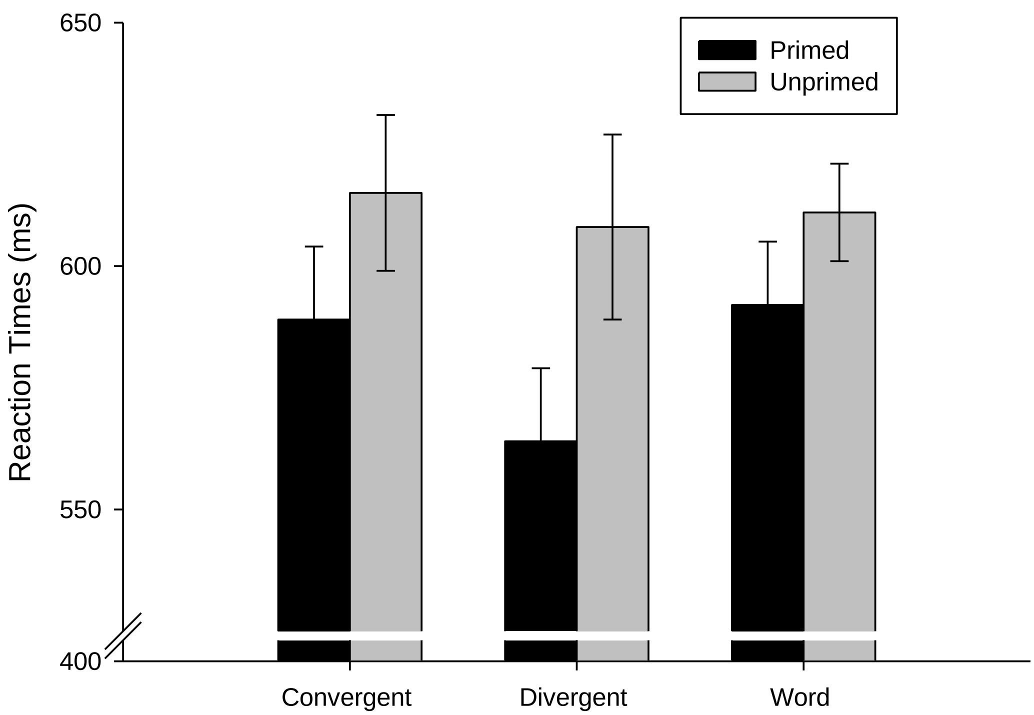

Figure 4.1. Priming effects for critical lures by condition.

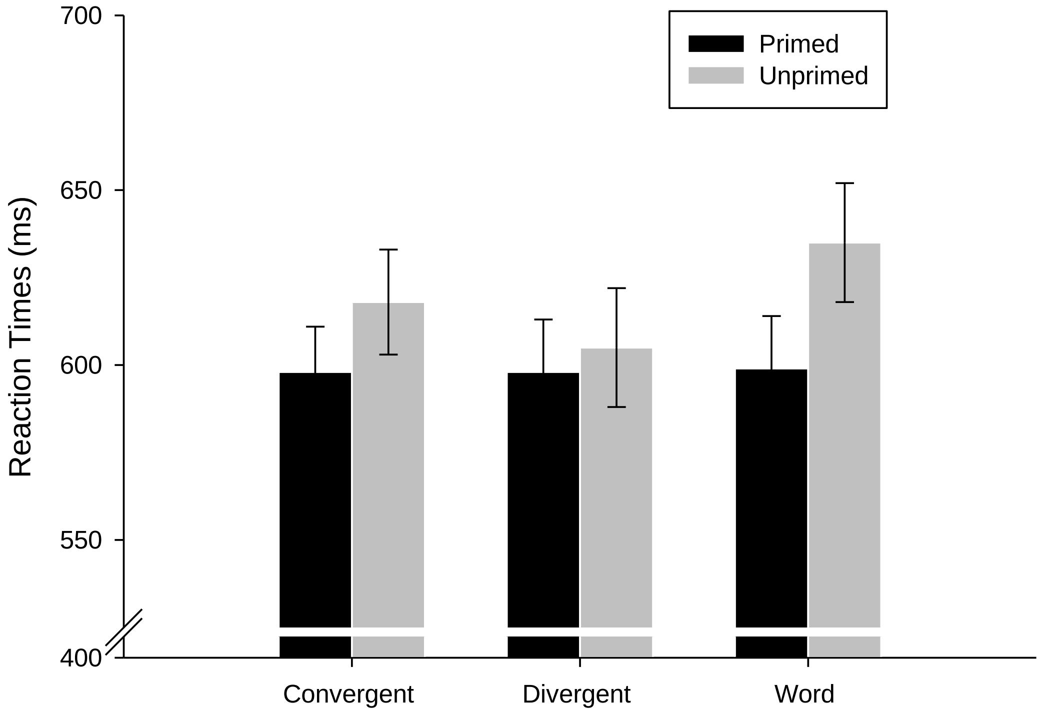

Figure 4.2. Priming effects for associates by condition.

Discussion

The purpose of Experiment 2B was to examine whether differences in critical lure activation between conditions could explain the differences in false alarm rates in Experiment 2A. The results suggest that in each condition critical lures primed by previous study lists were highly activated relative to control-matched unprimed items. Importantly, however, these priming effects did not differ across conditions and leaves the question unanswered as to why there were differences in false recognition between conditions during the recognition test of Experiment 2A.

Because the critical lure was at least initially internally activated in a similar manner across conditions, the differences in false recognition could be due to activation during retrieval.

That is, the critical lure may cue the episodic representations of studied items within the associative network that may activate the critical lure (Gallo, 2006). Thus, the critical lure may be *reactivated* at test (Meade et al., 2007) and this activation may be more consistent with the related studied items in the convergent and word-only conditions. Because relational processing was presumably disrupted in the divergent condition, even if the critical lure was reactivated at test, participants may be less likely to attribute this activation as a previously seen item since it is inconsistent with the studied items. This decreased (or inconsistent) reactivation hypothesis could account for both the lower veridical and false memories in the divergent condition.

CHAPTER 5

Experiment 3

In the divergent condition of Experiment 2A, items within-lists were unrelated to one another and did not elicit the meaning of the critical lure. Decreased false alarm rates could be due to either of these reasons. Therefore, we do not know whether it is within-list relationships, relationships between the associates and the lure, or some combination of both that is producing differential false alarm rates between conditions. Experiment 3 was designed to extend the results from Experiment 2A using associative items other than those typically used in the DRM procedure by placing associates in the context of sentences that converged on two separate meanings of a non-presented critical homographic lure. So, for example, associates were placed in the context of sentences that converged on each meaning of the word "fall" (e.g. "The flowers bloomed in the spring," and, "The slick ice caused her to slip."). In one condition, sentences from one meaning of the homographic lure were presented in succession and then sentences from the other meaning were presented. In another condition, presentation of sentences was alternated between the two meanings. These conditions were compared to similar conditions in which words were presented in isolation. Because associates similarly elicit the meaning of the critical lure in all conditions, differences in false recognition when presentation order is alternated would suggest that relational encoding of items within-lists may underlie the false memory effect. However, if no differences occur between successive and alternating presentation, it could suggest that associative convergence on any meaning of the critical lure is the primary factor in eliciting false recognition. Furthermore, we believe that sentence processing should result in greater (or at least equivalent) false alarm rates than word processing.

Methods

Participants. Undergraduate students from the University of Georgia volunteered in exchange for partial credit toward a course research requirement. Each participant was individually tested in sessions that lasted approximately 20 minutes. 132 new participants were randomly assigned to one of the four between-subjects encoding conditions (33 in each).

Materials. A total of 12 homograph lists were used from Hutchinson and Balota (2005) based off of the Twilley, Dixon, Taylor, and Clark (1994) norms. For the two separate meanings of the critical homograph, 4 associates were taken from each list. For example, for the homograph "fall", 4 words were related to the "autumn" meaning (e.g. "autumn", "season", "spring", "leaves") and 4 words were related to the "stumble" meaning (e.g. "stumble", "slip", "rise", "trip") to composed the 8-item list. The average backward associative strength (BAS) from each word to the critical homograph was equated between each meaning. We also created sentences for each of the 12 homograph lists, with 4 sentences related to one meaning (e.g. "The young boy hated raking leaves.”), and 4 sentences related to the other meaning (e.g. "The slick ice caused her to slip.") of the lure.

Design and procedure. The only difference between the two sentence conditions was the order of presentation. In the *sentence-grouped* condition, the 4 sentences from one meaning were presented in succession, and then the 4 sentences from the alternate meaning were presented in succession. In the *sentence-mixed* condition, the sentences from the two different meanings were presented in alternating fashion. The same structure followed for the two word conditions, however, the associates were presented in isolation (i.e. no sentences). In the *word-grouped* condition, the 4 associates from each meaning were presented in succession, and in the *word-mixed* condition presentation of the associates alternated between meanings.

The procedure used in Experiment 3 was nearly identical to Experiment 2A. However in the grouped conditions, the 4 sentences (or words) from one meaning were randomly presented and then the 4 from the alternate meaning were presented, counterbalanced across participants as to which meaning was presented first. In the mixed conditions, the stimuli from the two different meanings were randomly selected to be presented in alternating form. After making meaningfulness ratings on all 12 blocks, participants engaged in a 2-minute distractor phase and then were given instructions for the test phase. Instructions for the study and test phase in each condition were identical to those given in Experiment 1 and 2A.

The test phase consisted of 48 old and 48 new items randomly presented. In each condition, 2 old items were taken from each meaning of the studied list (resulting in 4 items per list). The new items in all conditions consisted of the 12 non-presented critical homographs, as well as 4 homographs taken from the norms that were never studied, with 3 associates from each list. There were also 16 unrelated new items taken from other homograph lists.

Results

Table 5.1 displays hit rates, false alarm rates to critical lures, and false alarm rates to unrelated lures. As with Experiment 2A, we employed a correction for hits and false alarms to critical lures due to differences in false alarms to unrelated lures across conditions ($F(3, 128) = 10.53$, $p < .001$, $\eta_p^2 = .198$). We conducted a 2 (item type: studied vs. critical lure) x 2 (context: sentence vs. word) x 2 (presentation: grouped vs. mixed) mixed ANOVA for average recognition. The analysis of corrected hit and false recognition scores revealed a main effect of item type, $F(1, 128) = 359.01$, $p < .001$, $\eta_p^2 = .737$, whereby critical lures (M = .40) were recognized less than studied items (M = .71). There was no main effect of presentation, $F(1, 128) = .003$, $p = .95$, $\eta_p^2 < .001$. A main effect of context was found, $F(1, 128) = 24.07$, $p < .001$, $\eta_p^2 = .158$, whereby more items were labeled "old" during word-only encoding (M = .60) than during sentence encoding (M = .51). There was also a significant interaction of item type and context, $F(1,128) = 160.75$, $p < .001$, $\eta_p^2 = .557$.

Table 5.1
Recognition Hit Rates and False Alarm Rates (Standard Errors) by Condition for Experiment 3

	Hits		False Alarms		
Condition	Studied	Corrected	CL	Corrected	Unrelated
Sentence-grouped	.75 (.02)	.55 (.02)	.68 (.03)	.48 (.02)	.20 (.02)
Sentence-mixed	.73 (.02)	.56 (.02)	.60 (.03)	.44 (.03)	.17 (.02)
Word-grouped	.94 (.01)	.85 (.02)	.42 (.04)	.34 (.03)	.09 (.02)
Word-mixed	.96 (.01)	.87 (.02)	.45 (.04)	.36 (.03)	.09 (.02)

Separate ANOVAs were conducted for corrected hit and false alarm rates. There was a significant difference in studied items recognized between the two encoding contexts, $F(1,130) = 188.83$, $p < .001$, $\eta_p^2 = .592$, whereby participants in the word-only (M = .86) conditions recognized more studied items than the sentence (M = .56) conditions. There was also a significant difference in false alarms to critical lures between conditions, $F(1,130) = 14.53$, $p < .001$, $\eta_p^2 = .101$. This comparison revealed that participants in the sentence (M = .46) conditions false alarmed to critical lures more often than participants in the word-only (M = .35) conditions. Thus, participants in the word-only conditions not only recognized more studied items than the sentence conditions but also false recognized fewer critical lures.

Discussion

The purpose of Experiment 3 was to examine whether differences in false alarms arose between sentence and word-only processing when stimuli converged on separate meanings of a homographic lure. As with Experiment 2A, we found that veridical recognition was better for word-only relative to sentence encoding conditions. In contrast, we found significant differences in false alarms to critical lures between sentence and word-only conditions that converged on the meaning of the critical lure, with greater false memories in the sentence conditions. Furthermore, our manipulation to reduce relational processing by alternating presentation of homographic meaning failed to produce any differences in hits or false alarm within the sentence conditions or word-only conditions.

Obviously, this experiment was different from the previous studies using stimuli other than those typically used in the DRM paradigm that diverge on two separate meanings of the critical lure. One possibility for the discrepancies in false alarms to critical lures between sentences and word-only conditions is that within each sentence, multiple pieces of information converge on the critical lure. For example, in the sentence, "The slick ice caused her to slip," the words, "slick," "ice," and "slip" could activate the critical lure "fall". Similarly, "The young boy hated raking leaves" has multiple pieces of information that activate "fall". Therefore, these rich semantic representations activate multiple associative pathways that converge on the critical lure. In the word-only conditions, only "slip" and "leaves" would activate the critical lure. Robinson and Roediger (1997) found that increasing the number of associates within lists increased the probability of false recall, and a similar mechanism could be influencing our results. Furthermore, this elaborative associative network may be more likely to be reactivated during retrieval thus making discrimination between internally generated and actually perceived items more difficult. This account could easily explain the differences in false alarms rates between the two types of encoding conditions. However, this does not necessarily explain why our manipulation to reduce relational processing by alternating presentation did not elicit differences in hits or false alarms. Because the stimuli converged on two separate meanings of the critical lure, relational processing may have been disrupted regardless of whether the presentation of homographic meaning alternated. Thus, as in Experiment 2A, diverging on the meaning of the critical lure reduced the production of shared cues that related sentences (words) together.

CHAPTER 6

General Discussion

Across three experiments we examined the influence of semantic context in the processing of DRM associates imbedded in sentences. Previous research suggests that semantic processing influences false recognition by strengthening semantic relationships among items making it more likely that the critical lure will be internally activated within an associative network (Toglia et al., 1999). However, making stimuli more distinctive by providing contextual information or encoding instructions that direct attention to differences among stimuli serves to reduce false memories (e.g. Goodwin et al., 2001; McCabe et al., 2004). The present study demonstrates that additional contextual information does not necessarily reduce false memories and can actually increase false memories in the DRM paradigm depending on the semantic properties of the stimuli. We provided contextual information that could be used to discriminate old from new items by placing DRM associates in sentences, finding that false memories were governed by the semantic properties of the stimuli that allowed for meaningful organization based on the similarities of the items. Experiment 1 demonstrated that the meaning elicited by encoded stimuli influenced false memories. That is, the more meaningful items were perceived (as indicated by subjective ratings) the greater the false alarm rates. Presumably, participants were able to form stronger relationships among items when processing allowed for more meaningful comprehension of the sentences. Experiments 2A and 3 extended these findings by suggesting that it is not simply how meaningfully the items are perceived that influences false memories (subjective ratings were equivalent across conditions), but rather the ability of the inherent properties of the stimuli to produce both relationships among studied items and connections from the items to the critical lure. Thus, organization of items in the DRM paradigm that allow for meaningful relational processing of items within-lists and that converge on the semantic meaning of the critical lure increases the likelihood that the critical lure will be activated (or reactivated) resulting in more source-monitoring errors.

In Experiment 2A, participants were much more likely to falsely accept a critical lure as old when sentences converged on the meaning of the lure. We believe this is because in the convergent and word-only conditions, not only are the stimuli related to other items within-lists (internal convergence), but they also converge on the meaning of the critical lure (external convergence). Internal and external convergence allow for meaningful organization of the items within-lists and form relationships from the item to the critical lure, respectively. In the divergent condition, however, stimuli lacked both internal and external convergence, which may have resulted in both less veridical and false recognition. Furthermore, in Experiment 1 meaningful blocks presumably had both internal and external convergence whereas either may have been disrupted in the meaningless blocks. That is, participants may have had difficulty either relating items together in the meaningless blocks or may have not been able to form strong associative connections between the sentences and the critical lure. However, veridical recognition was identical between blocks, which suggests that processing meaningless sentences may have disrupted external convergence. We therefore conclude from Experiments 1 and 2A that internal convergence improves veridical memory by increasing semantic relationships among items within a list while also increasing the probability that the critical lure will be implicitly activated. External convergence does not necessarily facilitate veridical recognition, but increases false recognition because stimuli elicit the meaning of the critical lure therefore increasing activation.

Furthermore, external convergence may be important during the retrieval process in order for the critical lure to cue the episodic representations of the studied items.

Using homographic lures in Experiment 3 we found slightly different results from that of Experiment 2A. Although different stimuli were used in the two experiments and cross-experimental comparisons are not definitive, we believe these differences warrant further discussion. While we replicated the differences in veridical recognition between sentence and word-only conditions, false recognition of critical lures was greater in the former than the latter in Experiment 3. Regardless of the fact that the homographic lure had two different meanings, both sentences and words similarly converged on one meaning of the lure (and they similarly converged on the other lure). Based on the previous experiment we should therefore not predict differences in false alarms between the two types of encoding conditions because both should have similar internal and external convergence. However, we argue that the differences arise due to the *degree* of external convergence (e.g. increased backward associative strength; Robinson & Roediger, 1997), whereby sentences create more elaborative associative connections between the stimuli and the critical lure. With this logic, however, it could be argued that similar increases in false alarms to critical lures should have been seen in the convergent relative to the word-only condition in Experiment 2A. However, because relational processing was not disrupted in the convergent or word-only conditions, the critical lure may have been activated due to the semantic relationships within-lists and the degree of external convergence was irrelevant. Because half of the items within-lists were related to one meaning of the homographic lure and half of the items were related to the other meaning, relating items together within-lists may have been more difficult (regardless of whether presentation of meaning was grouped or alternated). The differences in false recognition between sentence and word-only encoding conditions across experiments could be explained by the differences in internal convergence, with the degree of external convergence exerting a stronger effect when items within a list are more difficult to relate together. Regardless of the exact mechanisms involved, the results from the present study suggest that subjective organization imposed by the participant during encoding is influenced by the semantic context in which DRM associates are imbedded, and lists that are both internally and externally convergent increase lure activation and false alarm rates.

Relational Processing

Organization based on the commonalties of items is referred to as relational processing whereas processing the differences among items refers to item-specific processing (Hunt & Einstein, 1981). The item-specific/relational framework (Einstein & Hunt, 1980) suggests that the type of processing that occurs during encoding will influence memory performance at retrieval. Related list items should promote processing of shared information whereas unrelated lists should encourage processing of distinctive information (Hunt & McDaniel, 1993). Because divergent sentences (Experiment 2A) are arguably more distinctive than words, the item-specific/relational framework would suggest that item-specific information should be accessible in the divergent sentence condition and result in fewer false alarms to critical lures relative to encoding words in isolation. However, the lower veridical recognition in this condition (relative to the convergent condition) suggests that participants were unable to access item-specific information. Alternatively, the semantic orienting task in the current study may have encouraged item-specific processing in all conditions but related lists may have additionally allowed relational processing. The additive benefits of item-specific and relational processing in these conditions may account for the increased veridical and false memories. Using relational processing can increase veridical memory by responding "old" to items that are consistent with the overall theme of a studied list but also increase false recognition due to the similarities of the critical lures and studied items. In either case, we believe that processing of distinctive information in the current study resulted in a disruption of relational encoding by reducing internal convergence (within-list relationships) that decreased both veridical and false recognition. To further examine this account, future research could use free recall testing to examine organization of recall by measuring the adjusted ratio of clustering (ARC; Roenker, Thompson, & Brown, 1971), which is the extent to which participants cluster items according to category. Presumably, participants in the divergent sentence condition would have lower ARC scores than convergent conditions due to disrupted relational encoding. In addition, manipulating the orienting task (e.g. McCabe et al., 2004) to encourage relational processing may reduce the added benefits of relational processing in organized lists relative to unorganized lists, and a greater relative increase in false alarm rates may be seen in unorganized lists.

An experiment by Thomas & Sommers (2005) had participants read sentences that either converged on or diverged from the meaning of the critical lure or read words in isolation finding that younger adults we able to use the contextual information provided by the sentences to reduce false alarms to critical lures. However, we did not find definitive evidence that participants were able to access item-specific information to reduce false memories in Experiment 2A. Of course, there were several methodological differences between our study and theirs. First, we made their stimuli more homogenous by removing proper nouns and trying to equate the sentence length, which may have reduced the distinctiveness of items. They also had participants intentionally remember stimuli that may have allowed participants to focus more on item-specific information. Furthermore, the authors used a within-subjects manipulation in order to minimize the possibility that participants employed a "distinctiveness heuristic" whereas we used a between-subjects design. The distinctiveness heuristic suggests that when distinctive information (e.g. sentences versus words) is encoded between-subjects participants adopt a more conservative decision criteria to accept items as old due to a metamemorial expectation that more distinctive sentence information should be better remembered relative to word encoding (Dodson & Schacter, 2002; Schacter, Israel, & Racine, 1999; Israel & Schacter, 1997). Thus, the absence of sentence recollection would lead to the rejection of critical lures. However, when sentences and word lists are presented within-subjects, the distinctiveness heuristic is applied globally and similar false alarm rates occur for both classes of stimuli (McCabe et al., 2004). Although the distinctiveness heuristic could explain the reduction of false alarms in the divergent condition, it should also predict similar reductions in the convergent condition. It is unclear why participants would adopt differential decision criteria for rejecting lures in the two sentence conditions because both classes of stimuli are arguably more distinctive than the word-only condition. Rather, we believe that under incidental learning conditions participants organized information based on the similarities between the studied items and that the critical lure activation was more consistent with the studied items when there was both internal and external convergence.

Activation/Monitoring Theory

In terms of the activation/monitoring framework, encoding of the associates results in spreading activation to related items (i.e., the critical lure) within an associative network (Roediger et al., 2001). During subsequent recognition testing, participants may mistake internally activated representations with actual perceived items thus accepting critical lures as old. In Experiments 1 and 3 we are unable to determine whether differences in critical lure activation can account for differences in false recognition. However, the priming results from Experiment 2B suggest that the critical lure was equally activated across conditions.

Unfortunately, this rules out the possibility that differential false alarm rates across conditions are due to differences in initial activation. Furthermore, explaining spreading activation as reflected by priming effects in a LDT has been challenged (e.g. Brainerd, Yang, Reyna, Howe, & Mills, 2008) primarily due to the fact that semantic priming of the critical lure during a LDT is transitory whereas the DRM false memory effect persists over considerable delays (e.g. Seamon, Luo, Kopecky, Price, & Rothschild, 2002; Toglia et al., 1999). Although we found significant priming effects around 1 minute after the presentation of studied lists, these effects likely decay after the (roughly) 15 minutes between presentation of the first study list and recognition test in Experiment 2A. However, it is important to determine that this spreading activation does, at least initially, occur under different experimental conditions (and with incidental encoding instructions).

Interestingly, Thomas and Sommers (Experiment 3; 2005) had participants encode convergent (e.g. MATTRESS--BED) or divergent (e.g. RIVER--BED) paired associates and found similar results as our Experiment 2A. Participants were less likely to falsely recognize the critical lure (e.g. SLEEP) when the studied associates were presented in the context of divergent pairs. Furthermore, there were no differences in response latencies between the two conditions and no differences in latencies for hits and false alarms to critical lures. Finding no differences in RTs for veridical and false recognition suggests that in both conditions the critical lure was internally activated at test, but participants in the divergent condition were able to access item-specific information to reject critical lures. Although one could argue a similar process occurred in our study, Thomas and Sommers found no differences in veridical recognition between the two conditions. If the critical lure was similarly activated in the present study, we believe that decreased false recognition rates were due to inconsistencies between the activated lure and its relation to the studied items due to disrupted relational processing rather than increased item-specific processing.

Meade et al. (2007) had participants perform a LDT or a speeded recognition test after encoding a list of items and manipulated whether the critical lure appeared on either the first, third, sixth, or eleventh test trial after study. While priming effects occurred only when the critical lure was presented as the first item of the LDT, false recognition did not decline across test position. This suggests that although priming effects decay rapidly false recognition is persistent. They argued that when participants are placed in an episodic retrieval mode (Tulving, 1983) test items serve to *reactivate* the associative network and the degree of overlap between the items and the reactivated network facilitates recognition judgments. Upon presentation, the critical lure receives activation from the cued episodic representations within the associative network, thus subjects must use monitoring processes to discriminate between internal activation and activation from actual studied items. We similarly argue that when the critical lure is presented at test it reactivates the associative network, and conditions in which there is a high degree of external convergence (associative connections of studied items to the critical lure) there is a greater overlap between the critical lure and the reactivated network resulting in a greater number of source-monitoring errors. Although this hypothesis cannot be directly assessed in the current study, a recent study suggests that spreading activation can occur at test and increasing the number of tested associates (prior to presentation of the critical lure) increases critical lure false alarm rates for both studied and non-studied lists (Kimball et al., 2010). This suggests that when the critical lure was presented at test it overlapped considerably with the activated (or reactivated) associative network due to priming from tested associates, and that a spreading activation that occurs only during study is insufficient in explaining increased false alarms during recognition testing. To test the reactivation account, it would be interesting to manipulate the number of studied sentences in the convergent (e.g., four) versus divergent (e.g., eight) conditions to see if decreasing external convergence would result in a weaker reactivated associative network and reduce false memories in the convergent condition.

Alternative Theories

Although we have interpreted our results within the activation/monitoring framework, alternative theories have been proposed to explain the false memory effects in the DRM paradigm. As mentioned in the introduction, fuzzy-trace theory suggests that participants encode both verbatim and gist traces. When participants read meaningful sentences or encode stimuli that converge on the meaning of the critical lure, they may be more likely to extract gist representations of the critical lure due to increased semantic information related to the lure. To examine differences in gist extraction between conditions, inclusion instructions at test could be given for participants to accept both studied items and words related to studied items as old (Brainerd & Reyna, 1998). In addition, similar to our distinction between external and internal convergence, gist traces can be further separated into local and global gist (Neuschatz, Lampinen, Preston, Hawkins, & Toglia, 2002) where local gist reflects meaning extracted from individual items considered in isolation and global gist reflects the extraction of meaningful relations among items (Lampinen, Leding, Reed, & Odegard, 2006). In our study, false recognition could have been influenced by local gist extraction when stimuli converged on the meaning of the critical lure (external convergence) as well as through global gist extraction when stimuli were related within-lists (internal convergence). Alternatively, global matching models have also been used to explain false memories in the DRM paradigm (e.g. Hicks & Starns, 2006). Global matching models suggest that the target probe is matched to the features of all stored memory traces to produce a summed activation value. Critical lure errors occur because the lure has a small amount of similarities to stored memory traces of studied associates and the global match strength from the summed activation is sufficient to produce an "old" response (Arndt, 2010). This could account for the differences in false recognition in the current study by suggesting that the summed activation value was greater when sentences (or words) converged on the meaning of the critical lure due to the high degree of overlap between the test probe and stored memory traces. However, the interactive cueing hypothesis (Clark & Gronlund, 1996) proposes that items tested in the same context as studied should result in greater activation due to a match in both item and context information, which would suggest that false alarms in word-only encoding conditions should be greater. A more sensitive measure of this hypothesis would be to test studied sentences and new sentences that include critical lures. However, while both the fuzzy-trace theory and global matching models could explain our differences in false recognition across conditions, neither of these theories can readily explain the associative activation of the critical lure during the LDT (see Gallo, 2006).

Conclusion

In sum, the present study demonstrated the importance of contextual organization during encoding by showing that false memories are governed by semantic properties of the stimuli and the ability to activate an elaborate associative network. Although previous research suggests that distinctive processing can reduce the occurrence of false memories, this is not always the case. When associates were presented in the context of sentences, participants were more likely to falsely accept critical lures as old when the context allowed for meaningful relational processing of items. Internal convergence improves veridical recognition by enhancing relational processing that increases the semantic relationships among items and also increases the probability that the critical lure will be implicitly activated. External convergence influences false recognition by increasing lure activation due to the summation of activation produced by the multiple studied items. Furthermore, external convergence may also be important during the retrieval process in order for the critical lure to cue the episodic representations of the studied items, and to the degree in which an individual stimuli increases the associative connections, it may be more likely for the critical lure to be reactivated at retrieval. Thus, lists both high in internal and external convergence increase the probability that the critical lure will be activated and the considerable overlap between the critical lure and reactivated associative network increases monitoring failures. An important avenue for future research will be to empirically demonstrate the effects of varying semantic context on the reactivation process at retrieval and how this reactivation subsequently influences source-monitoring decisions.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, 22(3), 261-295.

Arndt, J. (2010). The role of memory activation in creating false memories of encoding context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 66-79.

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

Brainerd, C. J., & Reyna, V. F. (1998). When things that were never experienced are easier to 'remember' than things that were. *Psychological Science*, 9(6), 484-489.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11(5), 164-169.

Brainerd, C. J., Wright, R. R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(2), 307-327.

Brainerd, C. J., Yang, Y. Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in 'associative' false memory. *Psychonomic Bulletin & Review*, 15(6), 1035-10

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2(4), 331-350.

Bransford, J. D., & Franks, J. J. (1972). The abstraction of linguistic ideas: A review. *Cognition*, 1(2-3), 211-249.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, 3(1), 37-60.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, 11(6), 671-684.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58(1), 17-22.

Dewhurst, S. A., Pursglove, R. C., & Lewis, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5-year-olds. *Developmental Science*, 10(3), 374-378.

Dodd, M. D., & MacLeod, C. M. (2004). False recognition without intentional learning. *Psychonomic Bulletin & Review*, 11(1), 137-142.

Dodson, C. S., & Schacter, D. L. (2002). Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology and Aging*, 17(3), 405-415.

53.

Einstein, G. O., & Hunt, R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory*, 6(5), 588-598.

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York, NY US: Psychology Press.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, 38(7), 833-848.

Goodwin, K. A., Meissner, C. A., & Ericsson, K. (2001). Toward a model of false recall: Experimental manipulation of encoding context and the collection of verbal reports. *Memory & Cognition*, 29(6), 806-819.

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), 1-14.

Hege, A. G., & Dodson, C. S. (2004). Why Distinctive Information Reduces False Memories: Evidence for Both Impoverished Relational-Encoding and Distinctiveness Heuristic Accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(4), 787-795.

Hicks, J. L., & Marsh, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1195-1209.

Hicks, J. L., & Marsh, R. L. (2001). False recognition occurs more frequently during source identification than during old–new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(2), 375-383.

Hicks, J. L., & Starns, J. J. (2006). Remembering source evidence from associatively related items: explanations from a global matching model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1164-1173.

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), 1-28.

Hunt, R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior*, 20(5), 497-514.

Hunt, R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, 32(4), 421-445.

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

Johnson, M. K., Hashtroudi, S., & Lindsay, D. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3-28.

Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language*, 45(1), 82-102.

Kensinger, E. A., & Schacter, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology, 16,* 399–415.

Kimball, D. R., Muntean, W. J., & Smith, T. A. (2010). Dynamics of thematic activation in recognition testing. *Psychonomic Bulletin & Review*, 17(3), 355-361.

Lampinen, J., Leding, J. K., Reed, K. B., & Odegard, T. N. (2006). Global gist extraction in children and adults. *Memory*, 14(8), 952-964.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7(4), 560-572.

McCabe, D. P., Presmanes, A. G., Robertson, C. L., & Smith, A. D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review*, 11(6), 1074-1079.

Meade, M. L., Watson, J. M., Balota, D. A., & Roediger, H. (2007). The roles of spreading activation and retrieval mode in producing false recognition in the DRM paradigm. *Journal of Memory and Language*, 56(3), 305-320.

Neuschatz, J. S., Lampinen, J., Preston, E. L., Hawkins, E. R., & Toglia, M. P. (2002). The effect of memory schemata on memory and the phenomenological experience of naturalistic situations. *Applied Cognitive Psychology*, 16(6), 687-708.

Plancher, G., Nicolas, S., & Piolino, P. (2008). Influence of suggestion in the DRM paradigm: What state of consciousness is associated with false memory?. *Consciousness and Cognition: An International Journal*, 17(4), 1114-1122.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7(1), 1-75.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review*, 7(1), 158-162.

Robinson, K. J., & Roediger, H. (1997). Associative processes in false recall and false recognition. *Psychological Science*, 8(3), 231-237.

Roediger, H., & McDermott, K. B. (2004). In D. Balota, E. Marsh (Eds.) *Creating False Memories: Remembering Words Not Presented in Lists*. New York, NY US: Psychology Press.

Roediger, H., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, 8(3), 385-407.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, 76(1), 45-48.

Seamon, J. G., Lee, I. A., Toner, S. K., Wheeler, R. H., Goodkind, M. S., Birch, A. D. (2002). Thinking of critical words during study is unnecessary for false memory in the Deese, Roediger, and McDermott procedure. *Psychological Science*, 13(6), 526-31.

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(7), 1054-1064.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language, 40,* 1–24.

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(3), 424-432.

Thomas, A. K., & Sommers, M. S. (2005). Attention to item-specific processing eliminates age effects in false memories. *Journal of Memory and Language*, 52(1), 71-86.

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, 7(2), 233-256.

Tse, C., & Neely, J. H. (2005). Assessing activation without source monitoring in the DRM false memory paradigm. *Journal of Memory and Language*, 53(4), 532-550.

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

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

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, 70(1), 122-129.

Unsworth, N., & Brewer, G. A. (2010). Individual differences in false recall: A latent variable analysis. *Journal of Memory and Language*, 62(1), 19-34.