GENDER AND PREVIOUS EXPERIENCE INFLUENCE FALSE AND VERIDICAL MEMORY FOR DRM WORD LISTS IN HEALTHY YOUNG ADULTS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Leyla Linda Yousef

May 2012

ABSTRACT

Previous studies of the Deese-Roediger-McDermott (DRM) word list paradigm have not yet identified gender differences in the false recall of neutral critical lures, words with the strongest association to all words in a presented list. In the present study, male (n = 34) and female (n = 88) undergraduates studied DRM word lists for which "king," "sweet," and "window" were the critical lures. Following a distractor task, participants were given a written task of free recall. Based on their previous experience with word lists, participants were divided into "naïve" and "experienced" groups. Naïve females correctly recalled significantly more listed words (veridical recall) than males for the Sweet and Window lists, and were significantly less likely to produce the non-presented critical lure "window." No gender difference, however, was observed among the experienced group. Future DRM experiments should take into account the effects of DRM list characteristics as well as previous experience.

BIOGRAPHICAL SKETCH

Leyla Yousef graduated *cum laude* with a Bachelor's degree from Concordia University. During her time at Concordia University, she conducted research in the field of neuropharmacology. Following graduation, she pursued research in the field of molecular biology for several years. In May of 2012, she earned her Masters of Science degree in Nutritional Science with a minor in Cognitive Science from Cornell University in Ithaca, New York.

To my parents and grandmother

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to my Special Committee Chair, Dr. Virginia Utermohlen. I have been exceedingly fortunate to have you as my mentor. I am truly grateful for your countless hours of support, unwavering dedication, and guidance throughout this process. It has been a great honor and privilege to have been mentored by you.

I would like to thank the other members of my Special Committee; Dr. Barbara Strupp and Dr. Charles McCormick. Thank you for generously providing your time, support, and dedication in helping me attain my educational goals.

I would especially like to thank my family. This work would not have been possible without your love and support. Mom, there are not enough words to express how grateful I am to have you as my mother, thank you for being a devoted mom. Dad, your support and encouragement has always inspired me to do my best. Grandma, thank you for your guidance and love, you are my inspiration. Above all, I thank God for everything.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH III
DEDICATION IV
ACKNOWLEDGMENTS V
TABLE OF CONTENTS VI
LIST OF FIGURES IX
LIST OF TABLES X
CHAPTER 1 INTRODUCTION 1
Purpose 1
Overview 1
Definition of Key Terms 2
CHAPTER 2 LITERATURE REVIEW 5
False and Veridical Memory 5
Development of the DRM Paradigm 6
Construction of the DRM Lists 11
Variables Influencing False Memory 11
Word list variables 11
Critical lure variables 13
Theoretical Interpretations of the DRM Effect 15
Activation-monitoring theory 15
Associative activation 15
Source monitoring 18
Fuzzy-trace theory 20
Introduction to Short- and Long-term Memory 21
Memory as a Function of Individual Differences 23
Gender and episodic memory 24
Gender and spatial memory 26
Gender and the DRM paradigm 27
The Effect of Decay on False and Veridical Memory 30
Summation 32
Hypotheses 33

CHAPTER 3 METHOD **34**
PARTICIPANTS 34
MATERIALS AND DESIGN 35
PROCEDURE 36
DATA ANALYSES 39
  Data models 39
  Measures 40
CHAPTER 4 RESULTS **42**
VERIDICAL MEMORY PERFORMANCE 42
  Recall accuracy 42
  Data model: Overall significance 42
  Data model: Gender interaction 43
  Data model: Gender by DRM list type interaction 43
  Data model: DRM list type 45
  Items correct by DRM list type and gender 45
  Total items correct by experience 47
FALSE MEMORY PERFORMANCE 47
  Extra-list intrusions by DRM list type and gender 47
  Total extra-list intrusions by experience 48
  Total critical lures mentioned by gender 48
  Total critical lures mentioned by experience 49
  Critical lures mentioned by DRM list type and gender 49
VERIDICAL MEMORY AND MOOD STATES 51
  Items correct by mood states 51
  Extra-list intrusions by mood state, DRM list type, and gender 51
  Total critical lures mentioned by mood state 51
  Critical lures mentioned by mood state, DRM list type, and gender 51
ASSOCIATION OF VERIDICAL AND FALSE RECALL BY ATTENTIVENESS 52
CHAPTER 5 DISCUSSION **53**
VARIABLES INFLUENCING THE DRM EFFECT 53
  Previous exposure to word lists 53
  Attentiveness to memory tasks 55
  Self-reported mood states 55
GENDER INFLUENCES VERIDICAL AND FALSE RECALL 56
  Veridical recall 56
  False recall 57

Extra-list intrusions 58
Memory Performance Varies by DRM List Presented 59
Veridical recall 59
False recall 59
Limitations 62
Future Considerations 65
APPENDICES 67
Appendix A. Tables 67
Appendix B. Study Questionnaire 70
REFERENCES 86

LIST OF FIGURES

FIGURE 2.1. COMPONENTS OF SHORT- AND LONG-TERM MEMORY 23

FIGURE 4.1. GENDER DIFFERENCES IN VERIDICAL RECALL PERFORMANCE BY DRM LIST TYPE .. 46

FIGURE 4.2. GENDER DIFFERENCES IN FALSE RECALL PERFORMANCE BY DRM LIST TYPE 50

LIST OF TABLES

TABLE A1. THE THREE 15-WORD DEESE-ROEDIGER-MCDERMOTT (DRM) LISTS AND THEIR ASSOCIATED CRITICAL LURES USED IN THE PRESENT STUDY 67

TABLE A2. CHARACTERISTICS OF THE THREE 15-WORD DEESE-ROEDIGER-MCDERMOTT (DRM) LISTS USED IN THE PRESENT STUDY AND THEIR ASSOCIATED VALUES 68

TABLE A3. THE THREE 15-WORD DEESE-ROEDIGER-MCDERMOTT (DRM) LISTS WITH BACKWARD ASSOCIATIVE STRENGTH (BAS) AND INTER-ITEM ASSOCIATIVE STRENGTH VALUES USED IN THE PRESENT STUDY 69

CHAPTER 1

INTRODUCTION

Purpose

The DRM is a robust and often used technique in the study of false and veridical memory. Thus, it is of interest to understand if and how individual differences, such as gender, influence and contribute to the production of these memories in a healthy young adult population.

Numerous studies have determined that females excel in tasks of verbal and word list recall (Herlitz, Nilsson, & Backman, 1997). However, few studies have examined the influence of gender on the DRM as their primary focus. Understanding the effect of individual differences on the DRM, such as gender, will provide insight into false memory phenomenon and may be of importance to the design of future DRM experiments.

Overview

The present study examined whether gender influences false and veridical recall using DRM lists designed to elicit low and high rates of false recall. Male and female undergraduate students were presented with three lists, one eliciting a low rate of false recall and two eliciting a high rate of false recall. In addition, the latter two DRM lists also had distinct concreteness values.

We examined whether gender influenced the DRM task by employing a novel control and modifying previous methodologies. A unique feature of the DRM task is that it is usually administered to undergraduate university students. More specifically, students are commonly recruited by psychology departments and are usually offered extra credit for their participation. Many of these students may have either previously participated in DRM experiments or become familiar with the concept of false memory in a psychology course. This raises an interesting question: Does prior experience with word lists affect DRM performance? This exposure may influence the production of false memories, an individual difference that, to our knowledge, has not previously been implemented as a control.

Since there is a large amount of data supporting females' superior verbal and word recall performance, we hypothesize that females will exhibit superior performance on tasks of veridical recall. As such, the capacity of females' episodic memory would allow them to encode information to a greater extent than males. If they are better able to encode a memory, then one would expect that they would be able to suppress the production of false memories.

Definition of key terms

The following are definitions of key terms commonly found in the false memory literature that pertain to the present study. Words listed in parentheses are other possible terms for the same words commonly found in literature.

DRM word list: a list composed of 15 words that associate in a convergent manner to a non-presented word, termed the *critical lure*. These words are listed in descending order according to their associative strength. The DRM lists were originally adapted from word association norms (Nelson et al., 2004; Russell & Jenkins, 1954).

False memory: an erroneous representation of an item preserved in memory.

Veridical memory: an accurate or true representation of an item preserved in memory.

Extra-list intrusion (intrusion/unrelated intrusion/commission error): a non-presented word incorrectly recalled during free recall because it is semantically or phonologically associated to words on the DRM list. Due to a recency effect, words presented recently tend to intrude more than earlier words.

Critical lure (critical word): a particular type of extra-list intrusion, also not presented on the DRM list, with the strongest convergent associations to all the other presented words on the DRM list.

False recall (false memory/DRM illusion): the process of recalling critical lures from the DRM lists. False recall is measured as the number of words incorrectly recalled from the DRM word lists.

Veridical recall (correct recall): the process of recalling words correctly from the DRM lists.

CHAPTER 2

LITERATURE REVIEW

False and Veridical Memory

As we watch television, read, or speak with a friend, our mind is interpreting and integrating the content of these events based on the knowledge that we have acquired during our lifetime. This combination of experience and knowledge is known as an individual's schema (Bartlett, 1932). When memories are recalled, errors may be introduced as a result of the retrieved memory's assimilation with one's schema; these errors are known as *false memories*.

The Deese-Roediger-McDermott (DRM) paradigm provides a controlled context for studying the cognitive processes that underlie the creation of false memories. It has been thought that errors in memory may be introduced during the recall of meaningful memories or entire life events, but that they would not occur during the recall of something as simple as a word list. Interestingly, however, evidence of false memories has occurred in both instances (Roediger III & McDermott, 1995). This observation has led to an entire field of research investigating the production of false memories. The DRM is a standardized word list learning paradigm that was designed to elicit and quantify these memories. The characteristics of the DRM word lists themselves have become of interest because it is the nature of these word lists that affects the false memories produced. In addition to measuring the production of false memories, the DRM is also a measure of true or *veridical memory*.

Development of the DRM Paradigm

In 1894, E. A. Kirkpatrick stated that a person could sincerely report having the memory of an experience that did not actually occur as part of their experience. Furthermore, the same error in memory could even be created by more than one person. As Kirkpatrick spoke 10 common words to students, including "spool, thimble, and knife," they all at once thought of and wrote "thread, needle, and fork," respectively (Kirkpatrick, 1894). His observation was the first recognition of the false memory phenomenon.

Approximately 60 years later, (Deese, 1959b) began developing a process for creating a controlled experimental method to systematically produce errors in memory. At the time, Deese was not interested in false memory per se. He was more concerned with how these "intrusions" inadvertently affected his memory recall experiments. Deese hypothesized that he could predict the occurrence of these intrusions or false memories. An experimental task consisting of 36 word lists, 12 words per list, was presented to 50 undergraduate students. The lists Deese used were based on the highest frequency word associations to a stimulus word, termed the critical lure, from the Minnesota norms for the Kent-Rosanoff items (Russell

& Jenkins, 1954). He found that particular lists from the Minnesota norms resulted in participants generating these critical lures. The word-association norms collected by Russell and Jenkins were gathered from students in an introductory psychology course in the 1952-1953 academic year. This course was composed of 1,008 students, approximately 60% male and 40% female. The norms were based on students' free associative responses to stimulus words presented (Jenkins & Russell, 1960). To investigate the occurrence of false memories, students were instructed to listen to the lists and then orally recall the words. To obtain frequencies of word-associations, a free association word task was given to a separate group of students. Data gathered from both experimental groups revealed that the probability of an intrusion word occurring during a recall task is proportional to the frequency of the word's association to other words on the list. Similarly, Deese also found that the critical lure was recalled more frequently than other intrusions (Deese, 1959b). Thus, Deese identified word lists that produced the greatest number of false memories.

Thirty-six years later, Roediger III and McDermott (1995) built on Deese's work, using his task and six of the 36 word lists in their first experiment. They envisioned this experiment as an opportunity to study false memory using a simple task of free recall. In addition to the free recall task, they included a standard memory procedure, a recognition task. Roediger III and McDermott (1995) also developed six additional word lists consisting of 12 words per list. Word lists were presented orally at a rate of one word per 1.5 seconds to undergraduate students. Participants were asked to freely recall the words for 2.5 minutes and specifically instructed not to guess. Afterward, the recognition task was administered and students were instructed to rate the occurrence of words that had to be presented on a 4-point scale such as "sure old," "probably old," "probably new," and "new." Finally, words were read aloud and students were asked to raise their hands when critical lures were recognized.

During the recall task, words that were not presented on the lists were recalled at approximately the same rate as those that were actually presented. The critical lure was recalled with close to the same probability as those words presented in the middle of the list. Words recalled from the middle of the list are considered to exclude memory effects of primacy or recency. During the recognition task, recognition rates of non-presented critical lures were significantly greater than those of less related non-presented words. Critical lures were also reported as "sure old" or "probably old" at approximately the same rate as the presented studied words (Roediger III & McDermott, 1995).

In a second experiment, Roediger III and McDermott (1995) provided two new additions: word lists of slightly longer length and assessments of memory judgments. They developed 18 additional word lists, consisting of 15 words per list. During the recognition task, students were instructed to distinguish between "remembering" and "knowing" whether words they had written were observed on previous lists. "Remembering" entailed mentally rehearsing an experience, whereas "knowing" occurred when the participant was confident that the word was on the list but could not actually re-experience the memory. Critical lures of the new lists were recalled at a higher rate than those of the six lists from the previous study phase. In fact, critical lures were recalled at a higher rate than those of the presented studied words in the middle of the list. Critical lures were more likely to be judged as "remembered" rather than "known," indicating students remembered an experience that had never occurred (Roediger III & McDermott, 1995).

Following the Roediger III and McDermott (1995) experiment, the term "DRM paradigm," an acronym for Deese, Roediger III, and McDermott, was coined to describe the technique they had developed.

The present day DRM task consists of asking participants to read or listen to a DRM list of 15 words that have the strongest association to a missing word from the list, which is the critical lure (Nelson, McEvoy, & Schreiber, 2004; Russell & Jenkins, 1954). Participants are presented with words such as "bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn, and drowsy." All 15 words associatively converge on the critical non-presented word, "sleep." After words are presented, participants are instructed to freely recall as many of the words from the DRM list as they can remember, in no particular order. Depending on the experiment, participants may also be cautioned against guessing (Roediger III, Watson, McDermott, & Gallo, 2001).

What was once regarded as an extraneous nuisance has become a memory phenomenon of significant research interest. The DRM paradigm, in particular, is a commonly used and useful technique for investigating errors in memory. In fact, the six years following Roediger III and McDermott's 1995 article, one DRM experiment was published every two weeks (Roediger III, Watson et al., 2001).

There are several striking features of the DRM paradigm. Errors in memory occur upon immediate testing from a simple word list; no misleading information is necessary to elicit the errors; and errors occur even when participants are specifically instructed against making them (Roediger III, Watson et al., 2001).

Understanding false memory elicited by the DRM lists requires knowledge of how these lists are constructed, the variables that influence the differing rates of false memory, and the two main theories used to describe the phenomenon.

Construction of the DRM Lists

A remarkable feature of the DRM paradigm is that although all the word lists are constructed in a similar manner, they do not elicit the same rates of false recall. Thus, it is important to consider the underlying word list characteristics that produce this effect.

The 55 commonly used word lists are derived from the Roediger III, Watson et al. (2001) DRM lists and are composed of 15 words associated to a non-presented critical lure. These lists were originally adapted from word association norms (Nelson et al., 2004; Russell & Jenkins, 1954). DRM words are listed in descending order according to their associative strength.

Variables Influencing False Memory

A better understanding of the theories used to account for false memory first requires a brief description of the variables known to potentially influence the rates of false recall.

Word list variables. Word list variables include *inter-item associative strength, backward associative strength (BAS), forward associative strength (FAS),* and *veridical recall* (Cann, McRae, & Katz, 2011). Inter-item associative strength is the "average relative frequency with which all items in a list tend to elicit all other items in the same list as free associates" (Deese, 1959a, p. 305). Inter-item associative strength is positively correlated with veridical recall (Deese, 1959b; Roediger III, Watson et al., 2001). The greater the associations between DRM list words, the greater the veridical recall. Deese (1959a) also found that inter-item associative strength correlated negatively with the probability of false recall; Roediger III, Watson et al. (2001), however, found no relationship between the two. Furthermore, when extra-list intrusions occurred among lists with high inter-item associative strength, they were usually the same extra-list intrusions across participants (Deese, 1959a).

Mean backward associative strength (BAS) is the "average tendency for words in the study list to elicit the critical item on a free association test" (Roediger III, Watson et al., 2001). As words are presented, BAS either consciously or unconsciously activates the critical lure. Although BAS is considered an important variable in predicting veridical and false recall, it is apparent that there are other factors that elicit false memories. For example, the specific DRM list, known as King, has a high mean BAS value, but elicits low levels of false recall (Roediger III, Watson et al., 2001).

While BAS is a measure of the associative strength from the DRM list words to the critical lure, the mean forward associative strength (FAS) is a measure of the associative strength from the critical lure to the DRM list words (Roediger III, Watson et al., 2001).

Two DRM lists may have identical BAS and FAS values while eliciting very different rates of false recall. Therefore, there must be other contributing factors involved. After BAS (Deese, 1959b), the second strongest predictor is veridical recall (Roediger III, Watson et al., 2001). Veridical recall is the probability of correctly recalling DRM list words (Cann et al., 2011). BAS is positively correlated and veridical recall is negatively correlated with false recall. Accounting for 68% of the variance, BAS and veridical recall are the two main predictors of false recall (Gallo, 2006; Roediger III, Watson et al., 2001).

Critical lure variables. Critical lure variables include *word length*, *word frequency in language* (Kucera & Frances, 1967), and *rated concreteness* (Cann et al., 2011): that is, the degree to which words' referents are tangible objects. Obtained from the norms of Kucera and Frances (1967), the raw word frequency is the number of times the critical lure is found per million words in print. For example, "sleep" has a raw frequency of 65 and "butterfly" a raw frequency of 2. The concreteness rating for each critical lure is based on the word association norms of Nelson et al. (2004), which were originally obtained from the norms of Paivio, Yuille, and Madigan (1968) and the norms of Toglia and Battig (1978). For example, "sleep" has a midpoint rating of 4.74 and "butterfly" is more concrete with a rating of 5.91. The concreteness scale ranges from 1 to 7 (Roediger III, Watson et al., 2001), with 7 being the most concrete.

Although BAS is considered one of the most important determinants of false recall (Roediger III, Watson et al., 2001), the type of word associations may be just as important. There are three types of associative relations in memory: horizontal or coordinate (words linked at the same categorical levels); vertical or subordinate (words linked at different categorical levels); and proordinate or temporal (words linked in time or space) (Kihlstrom & Wilson, 1988). For example, the DRM list Window, which is considered a horizontal list, yields a high rate of false recall. On the same horizontal level, the activating items are "door" and "pane." In contrast, the Fruit list is considered a vertical list, yielding a low rate of false recall. On a subordinate level, the activating items are "apple" and "orange." In this case, the word "fruit" is a taxonomic category. Two experiments determined that the production of the critical lure is induced by lists with horizontal associations, and not by lists with vertical associations. Word lists of a basic categorical nature rarely produce false memories of a higher or superordinate nature (Buchanan, Brown, Cabeza, & Maitson, 1999; Pansky & Koriat, 2004). These experiments suggest that in addition to associative strength and direction, categorical structure is important as well. As in the case of subordinate categorization, one would not expect false recall of the word "needle" from the presentation of words from a lower (or more specific) category such as "hypodermic needle," "knitting needle," and "sewing needle." This research suggests that spreading activation occurs in a categorically horizontal manner in order to elicit false recall (Park, Shobe, & Kihlstrom, 2005).

Theoretical Interpretations of the DRM Effect

There are several cognitive processes at work involving the DRM effect. The two theories primarily discussed in the literature are the *activation-monitoring* and the *fuzzy-trace theor*y.

Activation-monitoring theory. The activation-monitoring theory describes two opposing processes involved during the DRM task. *Associative* a*ctivation* is the process responsible for activating the critical lure or contributing to the recall of a false memory. *Monitoring* is the process of cognitive editing that works to determine the origin of the activated memory. Associative activation promotes false memories while source monitoring reduces them (Gallo, 2010).

Associative activation. Associative activation involves spreading activation, which may occur during the presentation of the DRM list or during retrieval (Deese, 1959b; Roediger III, Balota, & Watson, 2001; Underwood, 1965). In contrast to associative activation theory, gist theory states that individuals create a mental gist representation during the presentation of the DRM list. This gist representation is a theme or summation of semantically related concepts that activates the critical lure due to its related nature

(Brainerd & Reyna, 2005). Gist theory is also articulated within the other primary false memory theory that will be addressed, the fuzzy-trace theory.

Spreading activation occurs when a concept is activated in episodic or semantic memory and spreads throughout an associative semantic network partially activating and thereby influencing implicit associative memory (Roediger III, Balota et al., 2001). For example, in a DRM list, the word "table" provides a greater false recall of the word "chair" compared to unrelated words such as "screen." The spreading activation theory dictates that while "table" was being encoded during presentation, it stimulated an implicit associative response to "chair" (Roediger III, Balota et al., 2001).

The strongest evidence in support of the associative-activation theory is the high correlation observed between false recall and BAS (Deese, 1959b; Gallo & Roediger III, 2002; Roediger III, Watson et al., 2001). The greater the associations of the list words to the critical lure, the more likely these words are to spread activation across the semantic network and elicit the critical lure. It has been suggested that any word list would most likely produce associative processing; however, it is only the lists with high BAS to the critical lure that actually elicit false memories (Deese, 1959b; Gallo & Roediger III, 2002; Roediger III, Watson et al., 2001).

Consistent with the idea of spreading activation, recall of critical lures increases as a function of DRM list length. A DRM experiment was conducted in which nine unrelated filler words were added to a list of 15 words. Veridical recall was reduced but there was no effect on the recall of critical lures. It was concluded that false recall is produced by the sum of the associative strength from words on the list (Murdock Jr., 1961).

One may discover how false memory is produced via spreading activation by investigating populations that have greater rates of false recall (Roediger III, Balota et al., 2001). Specifically, comparisons have been made between healthy young individuals and older adults; and between healthy older adults and those with Alzheimer's disease. Data from older individuals and those with cognitive impairments indicate intact automatic spreading activation mechanisms and impaired attentional processes (Balota, Black, & Cheney, 1992; Balota et al., 1999; Balota & Duchek, 1991). At least two pieces of information can be ascertained from this research. First, although older adults produce a greater rate of false memories than younger adults, older adults still maintain normal automatic spreading activation. This implies that the production of false memories is not due to greater levels of automatic spreading activation, but rather a failure to inhibit their creation. Second, the same inverse relationship pattern of false memory and veridical memory is observed in younger and older adults. Young individuals who have greater rates of false memory also have lower veridical recall (Roediger III, Watson et al., 2001). Similarly, older adults relative to younger adults have greater rates of false memory and lower veridical recall (Balota et al., 1999; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998).

Source monitoring. The second aspect of the activation-monitoring theory is source monitoring, an important mechanism contributing to individual differences observed in false memory production.

According to Unsworth and Brewer (2010), source monitoring is considered one of the primary predictors of false recall. Source monitoring is the cognitive process used to distinguish among all information brought to conscious awareness in order to make attributions about its origin (Johnson, Hashtroudi, & Stephen-Lindsay, 1993; Johnson & Raye, 1981). Source monitoring may occur during the encoding of a memory as well as during its retrieval. Sources from which information was originally perceived may include spatial, temporal, and social contexts (Johnson et al., 1993). For example, if one remembers that their colleague spoke about a particular fact but cannot remember on which of two occasions it occurred, then contextual discrimination could be made on the basis of a spatial-temporal detail. He or she may remember that their colleague was speaking in a particular lecture hall when the fact was stated. This concept of source monitoring will be discussed in further detail in a subsequent section.

Individuals with poor source monitoring capabilities are more likely to produce false memories. Unsworth and Brewer (2010) measured source monitoring ability using two source-monitoring tasks. In the first task of gender source recognition, participants were presented with new and old words in either a male or female voice. They were instructed to recognize the gender of the voice and whether the words were new or old. In the second task, picture source recognition, participants were presented with new and old pictures on a computer screen. They were instructed to indicate whether the pictures were new or old and in which of four distinct quadrants the old pictures were located. Structural equation modeling was used to determine whether source monitoring predicted false recall. It was found that individuals with poor source monitoring processing had greater rates of false recall than those with superior processing abilities. More specifically, there was a direct negative correlation of -0.58 between source monitoring ability and false recall. A strong positive correlation of 0.78 was obtained between source monitoring and veridical recall. After factoring out veridical recall, there was still a significant correlation of -0.36 between source monitoring and false recall. Therefore, source monitoring ability is a substantive shared variable in the relation between veridical recall and false recall (Unsworth & Brewer, 2010).

The production of false memories can be viewed as a failure of source monitoring, which results in one perceiving an experience as something it is not (Watson, McDermott, & Balota, 2004). Errors in source monitoring may occur for a variety of reasons, such as emotional self-focus or when one's attention is divided (Johnson et al., 1993). Studying the DRM word list activates a mental representation of the critical lure, which is an associated implicit response. If a word shares many similar contextual features with the presented words, then the word will be recalled. If it does not share these features, it will be edited, and thus not recalled. Therefore, a false memory error occurs when the source monitoring process mistakes the associated mental representation for the actual presented word and the editing process fails (Johnson & Raye, 1981; Watson et al., 2004). Individuals become unable to distinguish whether the critical lure has been internally or externally generated (Johnson et al., 1993).

Fuzzy-trace theory. According to the fuzzy-trace theory, encoding a memory requires two types of traces: a *gist trace* and a *verbatim trace*. While a gist trace captures the inherent meaning of an item or word, a verbatim trace corresponds to the precise representation of its surface form (Brainerd & Reyna, 2005). For example, one may have dissociated verbatim traces such as "drank a Sprite" and "ate turkey chili," and gist traces such as "drank a soda," "ate soup," and "had lunch." With regard to false memory, these two types of traces work in opposing directions. Gist traces support the recognition and recall of false memories such as "drank a Coke" and "ate chicken soup" while the retrieval of actual events (verbatim traces) suppresses their response through the process of *recollection rejection*. In this way, one may think "I did not drink a Coke or eat chicken soup because I clearly remember drinking a Sprite and eating turkey chili" (Brainerd & Wright, 2003; Brainerd, Reyna, & Mojardin, 1999).

As suggested by the above example, the recollection rejection is the process that occurs when there is a mismatch between a verbatim trace of a studied word and a mental association. For example, recollection rejection can occur when the verbatim trace and activated association have different orthographic and phonologic surface features or when individuals sense a different level of familiarity (Carneiro et al., 2012).

The applicability of the recollection rejection theory to the DRM paradigm has to do with its specificity. As opposed to other cognitive editing theories, recollection rejection operates at the item level (e.g., the word "Dallas" is rejected because of a verbatim mismatch with "Houston") rather than at the metacognitive level (e.g., a visual representation of an event distinguishing that only cities of Kansas were studied) (Brainerd, Wright, Reyna, & Mojardin, 2001).

Introduction to Short- and Long-term Memory

Figure 1 outlines the two basic forms of memory: *short-* and *long-term*. Short-term memory refers to remembering small amounts of information tested immediately or following a very short delay (Baddeley, Eysenck, &

Anderson, 2009). As implied, long-term memory is composed of the systems that store information over a long period of time.

Long-term memory includes two types of memory of interest: *semantic* and *episodic*. In 1972, Endel Tulving proposed a distinction between the two. He argued that semantic memory refers to general worldly knowledge, including sensory information such as taste and color (Tulving, 1972). Episodic memory refers to the remembrance of specific events. For example, when one hears that their best friend married, it becomes part of their semantic memory. Remembering when and where one learned the information becomes part of their episodic memory. Tulving (2002) emphasized that the term "episodic memory" only applies to memories that are re-experienced in some manner. For example, when sending a heartfelt thank-you note, one re-experiences a previous event in order to use that information to plan to write the note (Baddeley et al., 2009).

Short-term memory stores small amounts of information for a period of seconds to minutes. It is sometimes referred to as *working memory*; however, working memory is distinct in that it is a combination of storage and processing of complex tasks. Short-term memory can be further subdivided into *verbal* and *visual short-term memory*. Visual short-term memory includes visual and spatial short- term memory (Baddeley et al., 2009).

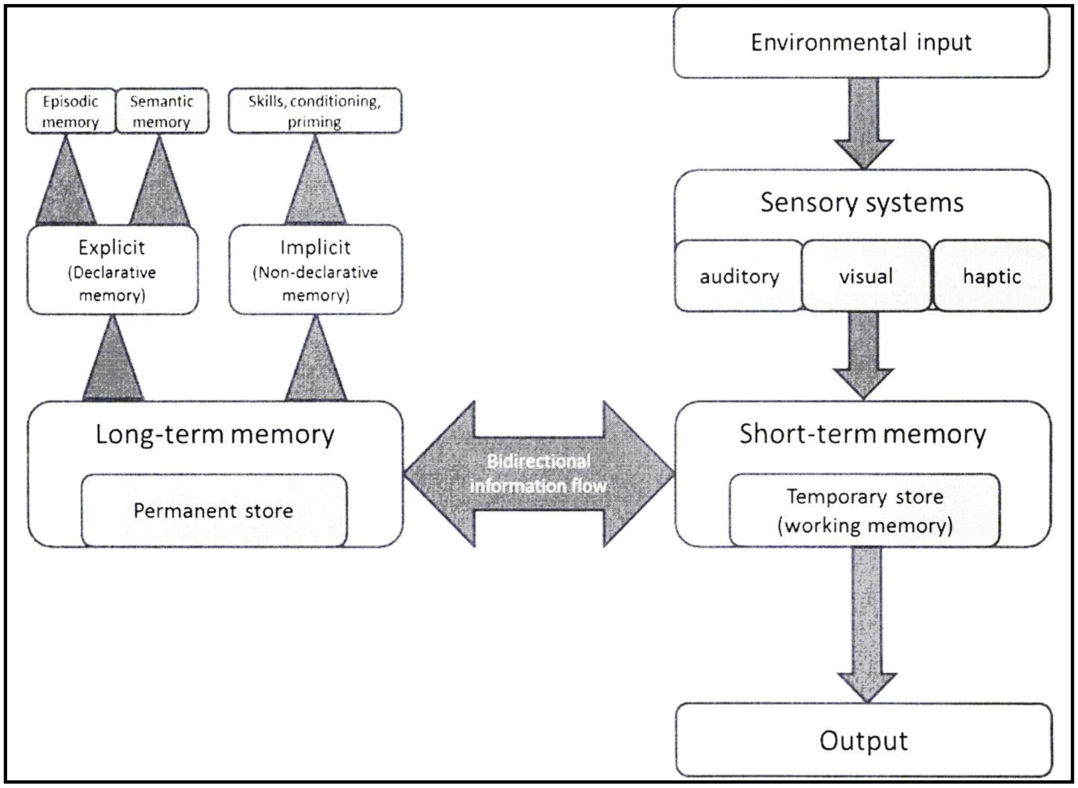


Figure 2.1. Short-term and long-term memory components, and their respective systems and processes (modified and adapted from Baddeley et al., 2009).

Memory as a Function of Individual Differences

False memory production also varies as a function of differences among individuals, such as gender, age, and cognitive function. This section will primarily address differences in episodic memory and false memory performance by gender. As previously mentioned, very few studies have been devoted to the investigation of the relationship between gender and false memory. Furthermore, gender differences and memory performance in general are usually observed as the by-product of another more central research question rather than as the sole purpose of the experiment.

Intrusions aside, the DRM can be more simply viewed as a task of word recall, composed of a list of associated words. Tasks such as these are used to measure episodic memory. Thus, it follows that in addition to measuring false memory production, the DRM task is also a measure of episodic memory, more commonly referred to in the false memory literature as veridical memory.

Gender and episodic memory. It is commonly observed that there is a female advantage on tasks of episodic memory, especially tasks of verbal memory. Furthermore, females' performance on these tasks is consistent among various encoding and retrieval conditions (Herlitz, Nilsson, & Backman, 1997). In the following three sets of experiments, females outperformed males on tasks of word recall. In a task of narrative recall, females also tended to reflect and evaluate information to a greater extent than males.

Hultsch, Masson, and Small (1991) investigated free recall in three different age groups. Five hundred and eighty-four participants from a metropolitan area were given a word recall test. The three age groups included: 19-36 years, 55-69 years, and 70-86 years. Six categorized lists from the Howard (1980) norms were visually presented. Participants studied these words for 2 minutes. Immediately following this presentation, they were given 5 minutes to freely recall the words.

Females recalled significantly more correct words than males (Hultsch et al., 1991).

Schaie and Willis (1993) also investigated free recall using word lists. Five hundred individuals, ages 20 to 70, were given an immediate recall test, a delayed recall test, and a word fluency test. During the immediate recall test, participants studied a list of 20 words for 3.5 minutes and were then given the same amount of time to freely recall the words. The same procedure was used in the delayed recall test, with the exception of the addition of a distractor period between the presentation and free recall. As a distractor, participants were given psychometric tests for 1 hour. During the word fluency test, participants freely recalled as many words as they could within 5 minutes. Females performed significantly better than males on all tests of verbal memory (Schaie & Willis, 1993).

Herlitz, Nilsson, and Backman (1997) conducted a large scale study investigating gender differences in four types of memory, including an episodic memory task involving word recall. Their population sample consisted of 1,000 participants, 35 to 80 years of age. Participants were auditorily presented with an uncategorized list of 12 common, unrelated words at 2-second intervals. Immediately following this presentation, participants were then given a test of free recall. Females recalled significantly more correct words than males (Herlitz et al., 1997).

Females create more detailed representations of past events and recall more information than males. In a study conducted by Bloise and Johnson (2007), undergraduates were asked to read narrative scripts and then given a written impromptu free recall test. Females recalled significantly more information than males when the information was neutral as well as when it was interspersed with emotional material. Participants were also asked to provide advice about the narrative scripts. Females gave more detailed factual advice than the males. This experiment suggests that females may recall, evaluate, and reflect upon information in a more detailed manner than males (Bloise & Johnson, 2007).

Gender and spatial memory. An evolutionary hypothesis, known as the *Hunter and Gathering Hypothesis*, argues that males predominantly hunted while females foraged or gathered. As an adaptation to these roles, females and males developed different spatial abilities. Males developed throwing and navigational skills. Females developed the ability to encode and recall spatial information of static objects, such as flora (Neave, Hamilton, Hutton, Tildesley, & Pickering, 2005). These two different types of spatial memory are known as *spatial navigational memory* and *object location memory* (King, Trinkler, Hartley, Vargha-Khadem, & Burgess, 2004; Postma, Kessels, & Van Asselen, 2008). Supportive of the Hunter and Gathering Hypothesis and contrary to the popular belief that males outperform females on all tasks involving spatial memory, many studies have shown that females outperform males on tasks of measuring object location memory (Honda & Yoshiaki, 2009; Voyer, Postma, Brake, & Imperato-McGinley, 2007).

Gender and the DRM paradigm. Given the findings on episodic memory, researchers have recognized that it is also worthwhile to investigate the influence of gender on the production of false memories.

In 2002, Seamon, Guerry, Marsh, and Tracy investigated whether gender significantly influences false recall. The sample size was composed of 50 males and 50 females who were given 16 DRM word lists with 15 words per list. Procedural instructions and words were given to each participant in the form of a spoken male voice at a rate of 1.5 seconds per word. No mention was made to the participants of the study's purpose or the DRM illusion effect. Immediately following the presentation, participants were asked to freely recall and write as many words as they could remember in 5 minutes. Although there were more critical lures produced (false recall) than correct words (veridical recall), this effect was not specific to a particular gender. Similarly, there was no significant gender difference observed for veridical recall, false recall, or extra-list intrusions (Seamon, Guerry, Marsh, & Tracy, 2002).

Bauste and Ferraro (2004) examined gender differences in false memory using DRM word lists, by presenting the list themes that were specific to the gender of the participant, such as MAN and GIRL. It was hypothesized that this gender specificity would invoke more false recall. The sample size consisted of 141 undergraduate students, 69 males and 72 females. Five DRM word lists (Man, Girl, Black, Bread, and High) containing 15 words per list were presented on a computer screen, one at a time for 1 minute. Following the fifth list, participants were asked to freely recall as many words as they could for 2 minutes. Following the recall task, in order to prevent rehearsal of word list content, a distractor was given in the form of a questionnaire relating to cognitive performance and health. Finally, participants were given a recognition task. There was no significant gender effect observed for the production of false memories under these conditions (Bauste & Ferraro, 2004).

Smeets, Jelicic, and Merckelbach (2006) conducted two studies to examine whether the effect of stress on false memory is influenced by gender. In the first study, the sample size consisted of 60 undergraduate students, 30 males and 30 females. Half the students were given a stress task: the TSST (Trier Social Stress Test); and half were given a filler task, such as a computer card game. Prior to and following the stress task or the filler task, a subjective mood state profile was obtained. Prior to, during, and following the DRM tasks, salivary cortisol samples were obtained. Twelve DRM word lists were used and verbally presented. Participants were then asked to freely recall the list words. During a recognition task, participants were presented with 72 studied and non-studied words. Males from the non-stressed control group produced significantly more intrusions than females. This effect was not observed in the stress-induced group. However, there was no evidence of gender influence for the proportion of correctly recalled presented words (veridical recall), falsely recalled critical lures (false recall), falsely recognized critical lures, or correctly recognized presented words.

In a second experiment, Smeets et al. (2006) increased the sample size to include 92 undergraduate students, 46 males and 46 females, using both low- and high-cortisol responders. The same experimental procedures were repeated with the exception of eight rather than 12 DRM lists, a 48-word recognition task, and modified allocation of the participant assignment groups; and the subjective mood state profile was obtained only once. As observed in the first experiment, males produced more extra-list intrusions than females; however, there was no observed gender effect for any of the other measures, indicating that stress did not affect false recall (Smeets et al., 2006).

Dewhurst, Anderson, and Knott (2012) examined whether gender differences would be observed in the false recall of emotionally negative DRM word lists. The sample size was composed of 50 females and 50 males who were given DRM lists containing neutral and negatively associated critical lures. Participants were shown 10 DRM lists on a computer screen for each type of negative and neutral content and then asked to freely recall the word lists following the study period. As part of the recall task, participants were instructed to write the 12 words from each list within 12 lined spaces. The mean BAS values for the lists were the same. Females correctly recalled a significantly greater number of words in general than males, with no specific list type effect. There was no significant gender effect observed in false recall; however, females recalled a greater number of false negative critical lures than males. Additionally, females produced a significantly greater number of false negative critical lures than neutral critical lures. Males recalled the negative DRM lists with significantly greater accuracy than females. However, there was no gender effect observed for neutral words (Dewhurst et al., 2012).

The Effect of Decay on False and Veridical Memory

Over time, activation of false and veridical memory within the semantic network is subject to varying rates of decay. This rate is dependent upon whether or not the individual is engaging in an *explicit retrieval strategy*, such as thinking back to the presentation of the DRM list. When individuals are not using this type of retrieval strategy and are using an *implicit activation strategy*, semantic priming experiments have demonstrated that the decay period of the critical lure is very short, lasting a maximum of several seconds

(Becker, Moscovitch, Behrmann, & Joordens, 1997; Dannenbring & Briand, 1982; Masson, 1995). *Semantic priming* refers to the increase in processing speed when a word is preceded by a related word (Meyer & Schvaneveldt, 1971). In accordance with the activation and monitoring framework theory, activation is spread more quickly among related concepts through the semantic memory network (Collins & Loftus, 1975). In contrast, when individuals use an explicit retrieval strategy, the decay period is much longer.

During tasks of false recall, false memory is less affected by decay than veridical memory (McDermott, 1996; Toglia, Neuschatz, & Goodwin, 1999), with false memories persisting up to 3 weeks following tasks of immediate free recall. Two studies in particular highlight the persistence of false memory. With the exception of the delay periods, Seamon, Chun et al., (2002) employed a typical DRM task that measured veridical and false recall over three retention periods: no delay, a 2-week delay, and a 2-month delay. Sixty undergraduate students were auditorily presented with a total of 16 DRM lists composed of 15 words per list, and were initially given a test of immediate free recall with no time limit. At each retention interval, participants recalled more critical lures than correct list words. While a 2-week delay produced a major decline in veridical recall, it had almost no effect on false recall. A 2-month delay produced a slightly further decrease in veridical recall but decreased false recall by about half (Seamon, Chun et al., 2002). A second and similar study by Toglia et al. (1999) observed that following immediate recall, false memories were just as strong 3 weeks following immediate recall.

Summation

The intent of the present study is to determine whether gender influences DRM performance. As stated in previous sections of this chapter, there is a vast amount of literature indicating that there is a female advantage for tasks of episodic memory; and more specifically, tasks of verbal and word recall (Herlitz et al., 1997). Since the DRM can be thought of as two types of tasks: measuring episodic memory (veridical memory) and false memory, it is likely that females would outperform males on the DRM for two reasons. First, the female advantage for verbal memory would allow females to better encode the information. Second, the literature has shown an inverse relationship between verbal memory and false memory (Roediger III, Watson et al., 2001). Thus, females would be better able to use their source monitoring processes to suppress the production of false memories.

Given the prevalence of DRM experiments conducted at universities, it is very likely that there would be a certain percentage of students that previously had exposure to DRM word lists. One would expect that these students' heightened level of awareness (source monitoring) for the nature of the task would enable them to perform better on the DRM task than those who did not have previous experience with word lists. Given that false recall slowly decays over time (McDermott, 1996; Toglia, Neuschatz, & Goodwin, 1999), the results of DRM performance would also depend on when the individual had previously performed a DRM task.

Hypotheses

We hypothesize the following:

1. Gender will influence performance on the DRM procedure; differences in performance between males and females will be observed in both veridical and false recall.
2. Among all participants, previous experience with word lists will improve overall performance on the DRM.

CHAPTER 3

METHOD

Participants

The complete sample consisted of 122 undergraduate students (*M* = 34 males, 88 females) from Cornell University. From this sample, 41% of students had previous experience with similar word lists. Therefore, the total sample was divided into a naïve group and an experienced group. The naïve group comprised 72 students (*n* = 21 males, 51 females) and the experienced group comprised 50 participants (*n* = 13 males, 37 females).

The ages of the participants volunteering in this experiment ranged from 19 to 28 years (*M* = 20.2, *SE* = 1.4). Seventy-seven percent of the participants were native English speakers and 23% were not. The distribution of their majors was 48% Human Biology, Health, and Society (HBHS); 38% Nutrition; 3% Human Development; 4% Biology and Society; 3% Biology; 1% Undeclared; 1% Animal Sciences; 1% Psychology; and 1% College Scholar. The experiment was conducted on May 2, 2007 during the class period of Human Anatomy and Physiology held in Martha Van Rensselear Hall, room 166. Students volunteered to participate in this experiment and were monitored by the teaching assistants in the class. This study was approved by the University Committee on Human Subjects (UCHS), a committee established to approve social and behavioral research conducted at Cornell University and its affiliates.

Materials and Design

The DRM word lists used in this study were adapted from Roediger III and McDermott's (1995) second experiment (see Table A1 in Appendix A), which were originally derived from the Russell and Jenkins' (1954) word association norms. Each list consisted of 15 words associated with a non-presented word, termed the critical lure. Three distinct DRM word lists were used and each list associatively converged on the critical lures: “king,” “sweet,” and “window.” Words presented in the King list included “queen, England, crown, prince, George, dictator, palace, throne, chess, rule, subjects, monarch, royal, leader, and reign.” The Sweet list included “sour, candy, sugar, bitter, good, taste, tooth, nice, honey, soda, chocolate, heart, cake, tart, and pie.” The Window list included “door, glass, pane, shade, ledge, sill, house, open, curtain, frame, view, breeze, sash, screen, and shutter.” These DRM lists were chosen for the following reasons: Word lists Sweet and Window are known to elicit the highest percentage of false recall: 54% and 65%, respectively. In contrast, the King list is known to elicit the lowest percentage of false recall, 10% (Stadler, Roediger III, & McDermott, 1999). Normative values, BAS, FAS, and concreteness for all DRM word lists used were included (see Table A2 in Appendix A and the section in Chapter 2 entitled "Variables Influencing False Memory").

Participants were given a questionnaire packet that included the three DRM lists, perception tests, self-assessed mood and health states, and demographic information (see Appendix B). DRM word lists were presented on separate pages with the 15 words, centered and double spaced. Each DRM list was followed by an exercise that served as a distractor in which participants were asked to describe what they thought was happening in each scene pictured on a separate page. In the DRM procedure, employing a distractor task circumvents a recency effect by preventing the rehearsal of study words in short-term memory until the period of free recall (Postman & Phillips, 1965). The results from the distractor task are not included in this study. Following the distractor, participants were asked to recall the words from the list on a blank page. Participants were given three different versions of the questionnaire, in which the lists and distractors were presented in different order. Additional information gathered from the questionnaire included health and demographic information and participants' self-assessed mood states.

Procedure

Students completed the DRM task and questionnaire during a regularly scheduled class period. They were verbally instructed not to open the test packets until asked to do so. One of three versions of the test was given to each student so that no two students seated next to each other received the same version.

Prior to opening the test materials, participants were verbally informed that the purpose of the task was to test their memory for word lists. When signaled, they opened their packets to the first page, which contained written instructions to study the first DRM word list for 30 seconds. When the 30 seconds had passed, a signal was given to move on to the next page, which displayed the distractor image. Participants were then asked to spend 2 minutes writing a story about what they thought was happening in the image. Following a signal that the 2 minutes had finished, they turned to the next page and were given 1 minute to write the words presented in the first list. This process (DRM list presentations, distractor image, and word recall) was repeated for each of the three DRM lists. In accordance with the Roediger III and McDermott (1995) paradigm, students were instructed not to guess during word list recall. Specifically, participants were instructed to freely recall and write down everything they felt confident was presented on the word lists. The final page of the DRM task included a question that asked whether participants had previous experience with word lists, similar to those used in this study. To ensure participants kept accurate timing and did not return to previous pages in their packets, they were monitored by the teaching assistants.

Participants' self-rated mood states were included as part of the questionnaire to determine whether their mood states affected their performance on the DRM. Students were asked to rate their mood states on the day of the DRM task by considering the percentage of time they had been "in a bad mood," "a little low or irritable," "in a mildly pleasant mood," and "a very good mood," totaling 100%. Mood states were split into sums of negative and positive mood states, totally 100%. Negative mood was defined as the percentage of time that they were in either a bad mood or an irritable mood. Positive mood was defined as the percentage of time that they were in an either pleasant or very good mood.

The questionnaire also included a task called the Ps and Qs Task, a neuropsychological test used to measure attention. In this study, the Ps and Qs Task was used as a control to confirm that both male and female participants performing the DRM task were engaged and attentive during the study. This task was also used to determine whether the attention, as a cognitive process, is a confounding variable in the DRM task.

Data Analyses

A secondary analysis (reanalysis from the original data collection) was conducted by the present researchers for the specific purpose of addressing the research question: How does gender influence the DRM effect?

Data were analyzed using the statistical data visualization software, JMP® 9.0.2, designed by the SAS Institute located at SAS Campus Drive, Building T, Cary, North Carolina 27513.

Data models. In order to conduct appropriate analyses, assumptions of normality were checked and data transformations were performed on non-normal or highly skewed data. For example, the number of words incorrectly recalled (extra-list intrusions) was converted from a continuous scale to a dichotomous scale (zero and one or more). When significant interactions were found within the statistical models, post-hoc tests were performed. Throughout this study, except where noted, all means were reported as plus or minus the standard error.

Analysis of Variance (ANOVA) models were performed on continuous data. Chi-squared analyses were performed on categorical data (e.g., a transformed extra-list intrusions variable). Regression analyses were performed to examine the relationship between continuous predictors and continuous outcomes.

Measures. Participants indicated whether they had prior experience with word lists; respectively, these groups were referred to as the *experienced group* and the *naïve group*. To control for this prior experience, separate statistical analyses were performed for the two groups.

As mentioned in Chapter 1, the primary measures of the DRM paradigm include veridical recall, false recall, extra-list intrusions, and the total number of critical lures mentioned. Veridical recall was defined as the number of presented words correctly recalled from the DRM study list. False recall was defined as the number of critical lures (words with the highest convergent association to all the other words on the list) produced that were not on the presented DRM list. Extra-list intrusions were words that were produced that were not presented on the DRM list (errors), including the critical lure. The total number of critical lures mentioned was calculated by taking the sum of critical lures collapsed across all three DRM lists.

Overall recall accuracy was calculated as the average recall accuracy across the three DRM lists. Plots of recall accuracy showed highly skewed distributions; therefore, the data was dichotomized by a median split. Chi-square analyses were used to examine differences in recall accuracy by gender.

To analyze self-rated mood states, mood states were split into sums of negative mood states and positive mood states, totaling 100%. Since the positive and negative mood states totaled 100%, it is statistically sufficient to only examine one of the mood states, as the other mood state would produce identical results. A median split was performed on the sum of the negative mood state variable. ANOVA models and chi-square analyses were used to examine the relationship between self-rated mood states and veridical recall, total critical lures, critical lures mentioned by gender, and extra-list intrusions produced.

CHAPTER 4

RESULTS

Veridical Memory Performance

Recall accuracy. Among the naïve group, the overall recall accuracy was 93.0%. A chi-square analysis showed no significant differences by gender.

Among the experienced group, the overall recall accuracy was 96.4%. A chi-square analysis showed no significant differences by gender.

Among all participants, the overall recall accuracy was 94.3%. A chi-square analysis showed no significant differences by gender.

Data model: Overall significance. Among the naïve group, an ANOVA model predicting the number of words correctly recalled using the independent variables gender, DRM list, and interaction of gender by list was significant ($F = 8.40$, $p < 0.0001$). The R^2 for the model was 0.168; therefore, the explanatory variables predicted 16.8% of the variation in the number of words correctly recalled.

Among the experienced group, the ANOVA model predicting the number of words correctly recalled using the independent variables gender, DRM list, and interaction of gender by DRM list was significant ($F = 4.71$, $p < 0.0005$). The R^2 for the model was 0.142; therefore, the explanatory variables predicted 14.2% of the variation in the number of words correctly recalled.

Among all participants, an ANOVA model predicting the number of words correctly recalled using the independent variables gender, DRM list, and interaction of gender by list was significant ($F = 13.03$, $p < 0.0001$). The R^2 for the model was 0.154; therefore, the explanatory variables predicted 15.4% of the variation in the number of words correctly recalled.

Data model: Gender interaction. Among the naïve group, averaged across the three lists, there was a trend toward significance in overall gender differences in veridical recall ($F = 3.38$, $p = 0.067$). Males averaged 9.52 ± 0.30 items correct per list, whereas females averaged 10.17 ± 0.19 items.

Among the experienced group, there was no trend observed, as with the naïve group, in overall gender differences in veridical memory ($F = 1.72$, $p = 0.192$). Males averaged 10.6 ± 0.5 items correct per list, whereas females averaged 10.6 ± 0.3 items correct per list.

Among all participants, averaged across the three lists, there was a significant overall gender differences in veridical recall ($F = 5.35$, $p = 0.021$). Males averaged 9.70 ± 0.2 items correct per list, whereas females averaged 10.31 ± 0.14 items correct per list.

Data model: Gender by DRM list type interaction. Among the naïve group, there was a significant gender by DRM list type interaction for veridical recall ($F = 3.14$, $p = 0.045$), indicating that the pattern of the females' performance was different from the pattern of the males' performance across the three DRM lists. Males recalled the greatest number of words from the King list (*M* = 11.05, *SE* = 0.51) and recalled the least number of words from the Window list (*M* = 7.76, *SE* = 0.51). Females recalled the greatest number of words from the Sweet list (*M*= 10.98, SE = 0.33) and recalled the least number of words from the Window list (*M* = 9.08, *SE* = 0.33).

Among the experienced group, there was a significant gender by DRM list type interaction for veridical recall (F = 0.19, p = 0.825), indicating that the pattern of the females' performance was different from the pattern of the males' performance across the three DRM lists. Males recalled the greatest number of words from the King list (*M* = 10.76, *SE* = 0.59) and recalled the least number of words from the Window list (*M* = 8.69, *SE* = 0.59). Females recalled the greatest number of words from the Sweet list (*M*= 11.19, SE = 0.36) and recalled the least number of words from the Window list (*M* = 9.43, *SE* = 0.35).

Among all participants, there was a significant gender by DRM list type interaction for veridical recall (F = 2.88, p = 0.057), indicating a trend that the pattern of the females' performance was different from the pattern of the males' performance across the three DRM lists. Males recalled the greatest number of words from the King list (*M* = 10.94, *SE* = 0.39) and recalled the least number of words from the Window list (*M* = 8.12, *SE* = 0.39). Females recalled the greatest number of words from the Sweet list (*M*= 11.07, SE =

0.24) and recalled the least number of words from the Window list (M = 9.22, SE = 0.24).

Data model: DRM list type. Among all naïve participants, independent of their gender, there is a significant difference between the lists in the number of words recalled correctly (F = 16.81, p < 0.0001). Students recalled the greatest number of words from the King list (M = 10.75, SE = 0.30) and the Sweet list (M = 10.37, SE = 0.31). The least number correct was recalled from the Window list (M = 8.42, SE= 0.30).

Among the experienced group, there is a significant difference between the lists in the number of words recalled correctly (F = 9.03, p < 0.0002). Students recalled the greatest number of words from the King list (M = 10.9, SE = 0.4) and the Sweet list (M = 10.9, SE = 0.4). The least number correct was recalled from the Window list (M = 9.1, SE = 0.4).

Among all participants, there is a significant difference between the lists in the number of words recalled correctly (F = 26.07, p < 0.0001). Students recalled the greatest number of words from the King list (M = 10.7, SE = 0.2) and the Sweet list (M = 10.8, SE = 0.2). The least number correct was recalled from the Window list (M = 8.9, SE = 0.2).

Items correct by DRM list type and gender. Among the naïve group, post-hoc tests from the ANOVA model showed that females recalled more words than males did from the Sweet list (F = 3.93, p = 0.049) and the

Window list (F = 4.69, p = 0.032), whereas there was no significant difference observed between the two genders in veridical recall for the King list (F = 0.96, p = 0.328) (Figure 4.1).

Among all participants, post-hoc tests from the ANOVA model showed that females recalled more words than males from the Sweet list (F = 4.76, p = 0.0298) and the Window list (F = 5.91, p = 0.016), whereas there was no significant difference observed between the two genders in veridical recall for the King list (F = 0.38, p = 0.537).

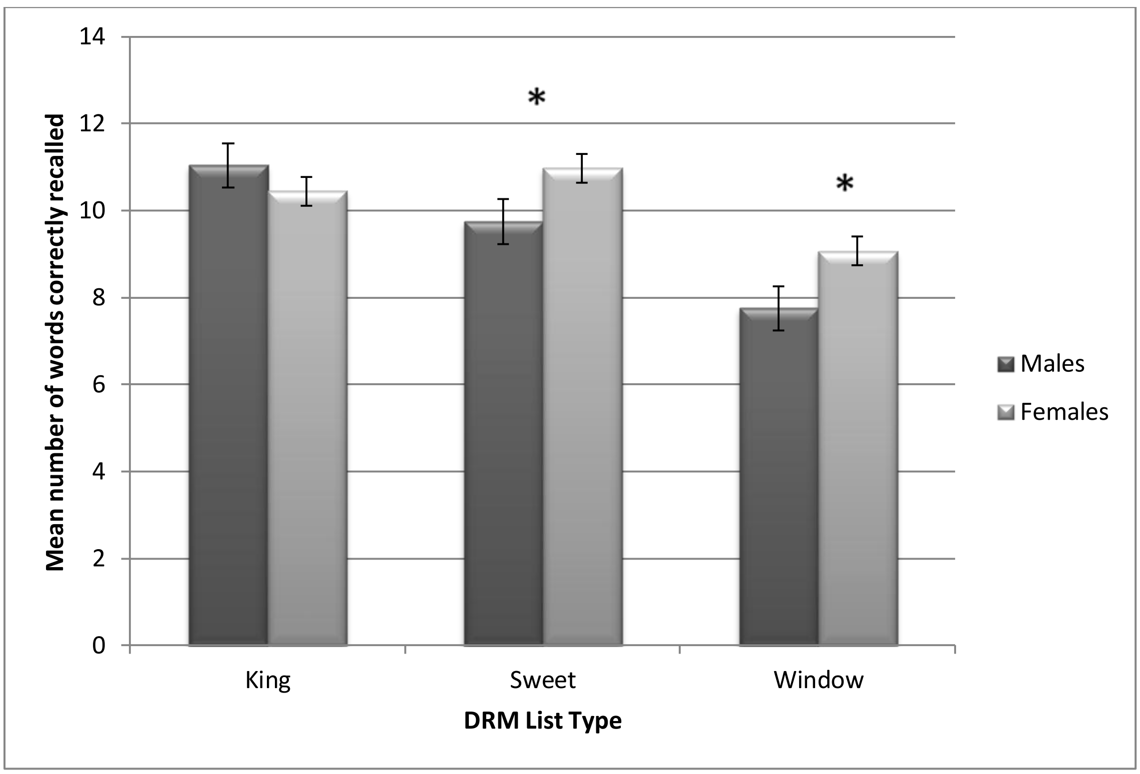


Figure 4.1. Gender differences in veridical recall performance by DRM list type. In the naïve group, females performed differently than males depending on the type of DRM list presented (F = 3.14, p = 0.045). Females recalled significantly more words correct than males from the Sweet list (F = 3.93, p = 0.049) and the Window list (F = 4.69, p = 0.032). No significant differences were observed in the King list.

Total items correct by experience. An ANOVA analysis indicated a borderline significant association (trend) between the naïve group and the experienced group (F = 3.44, p = 0.066) in total veridical recall performance across all three DRM lists. Individuals from the experienced group recalled slightly more items across the three lists than individuals from the naïve group.

False Memory Performance

Extra-list intrusions by DRM list type and gender. Among the naïve group, only 12 of the respondents gave more than one extra-list intrusion; therefore, the variable (number of extra-list intrusions) was transformed to a categorical variable with two levels: zero and one or more intrusions. The subsequent chi-square analysis comparing number of extra list intrusions between genders and between DRM lists showed no significant differences (all p > 0.05).

Among the experienced group, the chi-square analysis of the categorical number of extra-list intrusions also showed no significant differences by gender or by list (all p> 0.05).

Among all participants, the chi-square analysis of the categorical number of extra-list intrusions also showed no significant differences by gender or by list (all p> 0.05).

Total extra-list intrusions by experience. An ANOVA analysis indicated a significant difference between the naïve group and the experienced group (F= 4.67, p = 0.033) in the total production of extra-list intrusions across all three DRM lists. Individuals from the experienced group produced fewer incorrect items across the three lists than individuals from the naïve group.

Total critical lures mentioned by gender. Among the naïve group, the number of critical lures mentioned is defined as the sum of the three DRM list critical lures for each participant. A chi-square analysis of the total number of critical lures showed no significant difference by gender (chi-square = 3.44, p = 0.33). However, when the total number of critical lures mentioned was collapsed into two categories (one or fewer and two or more), a trend toward a gender difference was observed (chi-square = 3.22, p = 0.073). In males, 43% mentioned two or more critical lures, whereas in females only 22% mentioned two or more critical lures. There is an overall trend that males recalled more critical lures from the total of the three DRM lists than females.

Among the experienced group, a chi-square analysis of the total number of critical lures showed no significant difference by gender (chi-square = 4.33, p = 0.23). However, when the total number of critical lures mentioned was collapsed into two categories (one or fewer and two or more), a significant difference was observed (chi-square = 3.90, p = 0.049). Females recalled more critical lures from the total of the three DRM lists than males. In males, we saw that 0% recalled two or more critical lures, whereas 16.2% of females recalled two or more critical lures. The results of the chi-square analysis suggest an interaction between gender and experience. Therefore, a further analysis was conducted comparing the continuous variable (total number of critical lures mentioned) by gender and experience. The interaction of gender by experience was not significant (F = 2.5, p = 0.12).

Among all participants, a chi-square analysis of the total number of critical lures showed no significant difference by gender (chi-square = 0.75, p = 0.39).

Total critical lures mentioned by experience. A chi-square analysis indicated a significant difference between the naïve group and the experienced group in mentioning at least one critical lure across all three DRM lists (chi-square = 6.640, p = 0.010). Individuals from the experienced group mentioned significantly fewer critical lures (40.5%) than the naïve group (65.3%).

Critical lures mentioned by DRM list type and gender. Among the naïve group, a chi-square analysis of mentioning the critical lure "window" was significant by gender (chi-square = 3.92, p = 0.048). Therefore, a significantly greater percentage (66. 7%) of males produced the critical lure "window" than females (41.2%) (chi-square = 3.87, p = 0.049) (Figure 4.2). There was no significant difference observed for the King list (chi-square = 0.037, p = 0.848) or the Sweet list (chi-square = 0.37, p = 0.543).

Contrary to what was seen for the naïve group, a chi-square analysis of the critical lure recall by gender indicated no significant differences in any of the critical lures for the experienced group by list (all $p > .05$).

Among all participants, a chi-square analysis of the critical lure recall by gender indicated no significant differences in any of the critical lures by list (all $p > .05$).

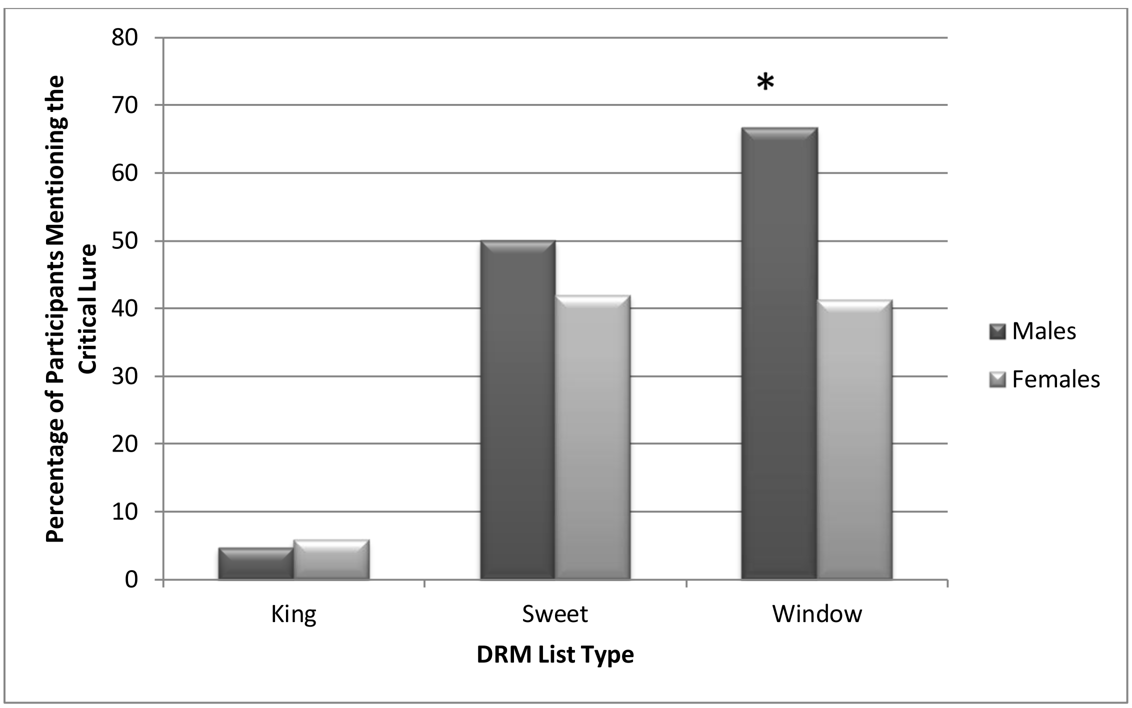


Figure 4.2. Gender differences in false recall performance by DRM list type. In the naïve group, chi-square analyses indicated that a greater percentage of males (66.67%) mentioned the critical lure "window" than females (41.18%) (chi-square = 3.92, p = 0.048). There were no significant differences observed for either the King list (chi-square = 0.037, p = 0.848) or the Sweet list (chi-square = 0.37, p = 0.543).

Veridical Memory and Mood States

An ANOVA model was used to examine the relationship between participants' self-rated mood states and veridical recall. Chi-square analyses were performed to examine the relationship between negative mood states and total critical lures mentioned, critical lures mentioned by gender, and extra-list intrusions produced.

Items correct by mood states. ANOVA analyses did not show any significant associations between mood states and veridical recall in any of the three lists. Results were similar for females and males in both the naïve and experienced groups (all $p > .05$).

Extra-list intrusions by mood state, DRM list type, and gender. Chi-square analyses did not show any significant association between mood states and the number of extra-list intrusions. Again, results were similar for females and males in both the naïve and experienced groups (all $p > .05$).

Total critical lures mentioned by mood state. Chi-square analyses did not show any significant association between mood states and the number of total critical lures mentioned. Again, results were similar in both the naïve and experienced groups (all $p > .05$).

Critical lures mentioned by mood state, DRM list type, and gender. Chi-square analyses did not show any significant association between mood states and the mention of any individual critical lure. Again, results were similar for females and males in both the naïve and experienced groups (all $p > .05$).

Association of Veridical and False Recall by Attentiveness

Among the naïve group, ANOVA analyses did not show any significant performance differences between males and females on the Ps and Qs Task, suggesting that males and females had similar levels of attention throughout the experiment. Chi-square analyses showed that performance on the Ps and Qs Task had no association with veridical or false recall (all $p > .05$).

CHAPTER 5

DISCUSSION

The purpose of this study was to determine whether individual differences, such as gender, influence veridical and false memories by using the Deese-Roediger-McDermott (DRM) paradigm. It is the first experiment to show that the influence of gender on memory performance is dependent upon the specific characteristics of the DRM list (see Table A2 in Appendix A).

Our experimental design has also brought to the surface an interesting confounding variable. It is common for undergraduate university students to be exposed either to the concept of false memory or to have performed DRM word list tasks. As a result of controlling for students' prior experience with word lists, we show for the first time that this exposure significantly impacts DRM task performance.

Variables Influencing the DRM Effect

Previous exposure to word lists. Due to the ubiquitous nature of DRM experiments conducted at universities, it is very likely that many students have previously performed the task and have learned the concepts relating to false memory in their psychology courses. We controlled for this by asking students to indicate whether they had prior experience with word lists at the end of the DRM task (see Appendix B). A surprisingly large proportion of students, 41%, indicated that they had. Our data indicate that previous experience with word lists does in fact lead to significantly different levels of performance on the DRM task, independent of the participants' gender. More specifically, those with experience outperformed individuals who were naïve to the task on measures of veridical recall; recalling more correct words across all DRM lists. Also across all lists, experienced participants produced significantly fewer intrusions and mentioned significantly fewer critical lures than did naïve participants. Our analyses also indicate that when individuals with previous experience are included in the total sample (as would normally be the case in DRM experiments), the observed false recall gender effect that was present in the naïve sample becomes eliminated.

The students that were exposed to previous word list tasks, such as the DRM, were most likely cognizant that one may produce errors in memory during word list tasks. As such, their increased awareness served to heighten their cognitive process, source monitoring. According to the activation-monitoring theory, source monitoring is negatively related to false recall (a correlation of -0.36) (Unsworth & Brewer, 2010). It is very likely that participants' enhanced source monitoring ability led to lower overall false recall and thus better DRM performance. Our data indicate that the gender differences are no longer present after controlling for previous experience. It is likely that males' heightened source monitoring processing brought their performance to the same level as females. These findings support the notion that previous experience improves performance on the DRM task and by not considering experience as a confounding variable, pertinent observations are neglected.

Attentiveness to memory tasks. One question that arises from the present findings is whether or not the overall poorer performance exhibited by males is an artifact due to a lack of attention or engagement with the DRM task. To control for this, participants were given a neuropsychological task called the Ps and Qs Task, which is a measure of attention (see Appendix B). Males and females performed equally well on this task, a finding which points out two items of relevance to the DRM results. First, that the males' overall poorer performance does not reflect a lack of attention or adherence to the experiment. Second, data also suggest that there is no significant correlation between performance on the Ps and Qs Task and performance on the DRM task, indicating that attention as measured by the Ps and Qs Task is not a predictor of false recall. In other words, in this experiment, attention as a cognitive process does not seem to influence DRM performance in either gender.

Self-reported mood states. A subjective mood state profile was obtained from participants as a control for the effect of stress on the DRM (Appendix B). No significant differences by mood were observed on any measure of the DRM performance, indicating that self-reported stress did not influence our results. However, it is worth noting that the one potentially emotionally charged list we used, “Sweet,” had both positive and negative words in it. We may have seen an effect had we chosen a more uniformly, negatively-charged DRM list, as was done by Dewhurst et al. (2012), which is discussed below.

Gender Influences Veridical and False Recall

Veridical recall. In the literature, it is commonly found that females outperform males on episodic memory tasks (Herlitz et al., 1997). In particular, there is a female advantage for verbal memory and, more specifically, tasks of word recall (Dewhurst et al., 2012; Herlitz et al., 1997; Hultsch et al., 1991; Schaie & Willis, 1993). The DRM task is a variant of word recall and because we provided a distractor between memorizing and recalling the word lists (see Appendix B), it was also a reflection of long-term episodic memory. Among the two lists known to elicit high rates of false recall (see Table A2), we found that females correctly recalled more presented words than males. Of the two studies that investigated veridical recall using the DRM task, neither demonstrated a gender difference. However, no distinction was made among individual DRM word lists in either of these studies (Seamon et al., 2002; Smeets et al., 2006). It is likely that a more detailed analysis of the data obtained in these studies would yield gender differences.

False recall. To the best of our knowledge, of the four studies investigating the influence of gender on the DRM effect (Bauste & Ferraro, 2004; Dewhurst et al., 2012; Seamon et al., 2002; Smeets et al., 2006), this study is the first to find a difference between males and females in the production of critical lures using neutral DRM word lists. Dewhurst et al. (2012), who used emotionally negative word lists, observed that females had higher rates of false recall when using the negative word lists but not when using neutral lists. In the present study, a significant difference was found when using the DRM Window list, an emotionally neutral list. By contrast, the Sweet list may contain emotional content but there was no difference by gender in critical lure production.

It is very likely that the gender effect was found because the approach used in the present study has allowed for a more detailed and therefore more discriminating representation of the data. It is commonly expressed in the literature that despite the fact that DRM lists are constructed in a similar manner, individually they elicit very different levels of false recall (Roediger III, Watson et al., 2001). Thus, it is important to consider the characteristics of each DRM list (see Table A2). By analyzing and interpreting each DRM list type separately, we were able to observe differences that would not have otherwise become apparent if the data were collapsed across the lists.

Extra-list intrusions. A critical lure can be considered a particular type of extra-list intrusion, one that is highly associated to all words on the DRM list and thus falsely recalled with the highest probability (Unsworth & Brewer, 2010). On the other hand, other extra-list intrusions produced may not be related to all the words on the list, suggesting weaker associations.

Although we found a gender difference in the number of critical lures produced, we did not find a difference in the production of extra-list intrusions. Thus, gender did not lead to the production of extra-list intrusions with weaker inter-item associative strength. Consistent with this idea, the majority of studies investigating the influence of gender on the DRM have also found no difference in the production of extra-list intrusions (Dewhurst et al., 2012; Seamon et al., 2002; but not Smeets et al., 2006).

The difference between the production rate of extra-list intrusions and of critical lures by gender leads to the speculation (discussed more fully below) concerning the production of critical lures, namely that males used a strategy different from that of females in order to remember the items on the Window list.

Memory Performance Varies by DRM List Presented

Veridical recall. In this study, females outperformed males on the task of veridical recall using the high false recall lists, but not the King list. An explanation for this observation may be that females were better able to cope with the increased processing demands of the Sweet and Window lists as compared to the King list. Since females are known to have superior episodic memory (Herlitz et al., 1997), particularly with word recall, they encode the verbatim trace to a greater extent than males. As a result, when source monitoring demands are higher (Meyers-Levy & Maheswaran, 1991), as is the case with lists known to elicit high rates of false recall (high gist), they are better able to suppress the production of critical lures.

False recall. As noted earlier, Dewhurst et al. (2012) found a gender difference using DRM word lists of a particular type, namely those that create an emotionally negative gist representation. This suggests that gender differences in false recall performance may depend on the gist elicited by the specific DRM list. BAS is the average tendency for words in the presented DRM list to elicit the critical item on a free association test (Roediger III, Watson et al., 2001). All the DRM lists used in this study had similar BAS values (see Tables A2 and A3 in Appendix A) (Nelson, McEvoy, & Schreiber, 1998). However, they differed significantly in concreteness (the rated tangibility of a critical lure based on word association norms) (see Table A2 in

Appendix A) (Nelson et al., 1998). The Window list has a higher concreteness value (6.27) than the Sweet list (4.53), with the King list concreteness value (5.54) intermediate between the two. It is with the more concrete list, Window, that we found a gender difference.

Highly concrete words can be represented in memory as an image as well as in the form of words. Altarriba, Bauer, and Benvenuto (1999) suggest that this dual representation helps with remembering concrete words, as the image provides an additional means of storage and retrieval. However, the drawback of concreteness is that the imaginal representation developed may contain extraneous items that are in some way connected to the words of the list. In the particular case of the Window list, many of the words contained in it, such as drapes, curtains, and shutters, represent objects that surround a physical window. Therefore, if one used a visual approach when encoding the list, as males may be inclined to use, given their gist-based processing and visuo-spatial advantage (King, Trinkler, Hartley, Vargha-Khadem, & Burgess, 2004; Postma, Kessels, & Van Asselen, 2008), a window image is highly likely to be visualized.

Why males produced more critical lures for a word high in concreteness may have to do with how males and females process gist-based information. When items to be remembered are associated, as is the case in DRM lists, it appears that females will tend to use a detail-oriented approach while males may take a more gist-dependent approach (Meyers-Levy & Maheswaran, 1991).

The gender effect may also be explained by the DRM lists' distribution of inter-item associative strength (see Table A3 in Appendix A), whose values were derived from students' free associative responses to the critical lures. The King list has the highest average inter-item associative strength value, followed by the Sweet list, and then by the Window list. Representing inter-item associative strength as an average, however, is misleading because it does not take into account the skewed distribution of values for each list. For example, the first item on King list is "queen," a word with a very high inter-item (free associative) value to the word King. Most of the remaining words on the list, however, do not have high associative values to the critical lure, "king." We suggest that because of the high inter-item associative value of "queen" with "king," an individual studying the list will easily recognize that the word "king" is missing from the list. As source monitoring is heightened, the individual will be enabled to reject the critical lure.

In contrast, the Window list contains items more evenly distributed in terms of inter-item associative value and no outlier is present (as is the case for the King list), thus lowering the list's average. Since the positive relationship of an item's inter-item associative strength (free associative) value to the critical lure is indicative of the propensity for spreading activation to occur through the semantic network, this may explain why males, who may be more inclined to use a gist-based approach, produced more critical lures than females for the Window list.

It is unclear whether the Window list elicited more critical lures for males because of its high concreteness or because of its more distributed inter-item associative strength (free association) values, or both; but what can be concluded is that performance differences on the DRM task are dependent on the specific characteristics of the DRM list used. Individual differences in DRM performance may also arise from the approach they typically use to memorize word lists.

Limitations

The present study has several limitations. First, a limited number of DRM lists were used. While studies investigating the influence of gender included as few as five DRM lists (Bauste & Ferraro, 2004), other studies have used as many as 20 lists. Using a greater number of lists would have allowed for a more in-depth comparison of various DRM list characteristics, such as the lists of extremely low/high concreteness values or low/high BAS values.

Second, the population sample is restricted by the composition of students enrolled in the Human Anatomy and Physiology course in which the experiment was conducted. The majority of students in this course majored in either Human Biology, Health, and Society (HBHS), or Nutrition, in contrast to previous studies of the DRM effect, in which the college students were psychology majors. It is possible that the effects ascertained in this study are a consequence of the characteristics of the population that participated, and that these results may not be found in other populations. In future experiments, it would be useful to control for students that had learned concepts relating to the DRM paradigm in their undergraduate courses, as well as to include students with majors unrelated to psychology, nutrition, or biology. Furthermore, it is necessary to conduct experiments in young adults not attending college, and use a different proportion of ethnic groups and other age groups.

Third, we did not include gender stereotypical DRM lists such as Man and Girl. Although Bauste and Ferraro (2004) did not observe a gender difference using these lists, it would be interesting to determine whether biology students may have different associations to the items in these lists, by virtue of having a greater interest in and understanding of the anatomical and physiological distinctions between males and females.

Fourth, we did not use a recognition task in addition to using a word recall task. Although word recall is known to be a more sensitive measure of the DRM effect than recognition (Hege & Dodson, 2004) and has been used in every gender study investigating the DRM paradigm, the addition of a recognition task would have provided a means for isolating the effects of source monitoring (Tse & Neely, 2007).

Fifth, because the question posed at the end of the DRM task was open-ended, students' interpretations may have varied. The question, "Have you had experience with word lists similar to the ones used in this study?" was phrased in such a manner as to capture the maximum number of individuals in the experienced group. Participants included in the experienced group may have interpreted the question more generally, without reference to the DRM task. It is also possible the students who indicated previous experience with word lists were more academically motivated and thus would perform better on tasks requiring the memorization of word lists. Without further information concerning the naïve and experienced groups, we cannot conclusively exclude these possibilities.

It is also important to keep in mind that one cannot interpret data from the DRM task to refer to "false memory" as it applies to the fabrication of entire events. Although there are some principles from the DRM paradigm that can be applied to inform our understanding of memory, word list tasks differ from entire autobiographical memories in personal relevance, emotional significance, and perceptual details. These differences limit any broad interpretations or generalizations from the DRM task to autobiographical memories (Gallo, 2010).

Future Considerations

The findings from this research spur many questions remaining to be answered. Since it was found that males and females perform differently on the DRM using a list high in concreteness and rate of false recall, how do males' and females' source monitoring processes differ to produce this effect? In conjunction with the suggestions made in this study, pairing the DRM experiment with a source monitoring task (Unsworth & Brewer, 2010) may provide insight to this question. Could a gender difference be found using lists with very high concreteness values (greater than 6) but low in false recall rate? Such a study may further promote an understanding of how males and females differ in their approach to processing gist-based information.

Broadly speaking, what has been learned from this study may be applicable to most DRM experiments in three ways. First, false memory researchers may be able to reinterpret their previous data in a novel manner. Reanalysis of the data from previous DRM experiments can be accomplished by parsing the data by DRM list type, such as low/high concreteness or low/high rates of false recall, as well as by gender. Second, future research can be designed taking into account additional important variables, such as participants' previous experience with word lists. Experiments can include a question at the end of the DRM task asking participants to indicate whether they had previous experience with word lists, and if so, what type and when.

It would also be helpful to know whether participants had previously learned the specific word lists being tested. Third, given the prevalence of DRM experiments conducted at universities and the significant impact that experience with word lists had on our results, it is highly recommended that future DRM experiments take the effect of experience into account. With regard to investigators interested in individual differences, such as gender or age comparisons, grouping and analyzing DRM lists by variable type may provide additional information as to how individuals process veridical and false memories differently.

In conclusion, it is of statistical and methodological importance to the scientific community using the DRM task to control for previous exposure to word lists; to group and perform separate analyses for DRM lists by type; and to take into account individual differences, such as gender.

APPENDICES

APPENDIX A. Tables

Table A1

The Three 15-Word Deese-Roediger-McDermott (DRM) Lists and Their Associated Critical Lures Used in the Present Study

King List	Sweet List	Window List
queen	sour	door
England	candy	glass
crown	sugar	pane
prince	bitter	shade
George	good	ledge
dictator	taste	sill
palace	tooth	house
throne	nice	open
chess	honey	curtain
rule	soda	frame
subjects	chocolate	view
monarch	heart	breeze
royal	cake	sash
leader	tart	screen
reign	pie	shutter

Note. DRM word lists were adapted from Roediger III and McDermott's (1995) second experiment developed from word association norms (Russell & Jenkins, 1954). Words are listed in descending order according to their associative strength.

Table A2

Characteristics of the Three 15-Word Deese-Roediger-McDermott (DRM) Lists Used in the Present Study and Their Associated Values

Word List	**False Recall**	**Veridical Recall**	**Concreteness**	**BAS**	**FAS**
King	0.10	0.650	5.54	0.230	0.059
Sweet	0.54	0.630	4.53	0.172	0.054
Window	0.65	0.630	6.27	0.184	0.058
Range	0.10 - 0.65	0.500 - 0.720	2.18 – 6.83	0.002 – 0.431	0.014 – 0.063

Note. Norm values were derived from Roediger III, Watson et al. (2001) and Stadler et al. (1999). BAS and FAS values were based on the associative norms from Nelson, McEvoy, and Schreiber (1998). False recall and veridical values were obtained from Roediger III, Watson et al. (2001). Concreteness values are based on the word association norms of Nelson et al. (2004), which are originally obtained from the norms of Paivio, Yuille, and Madigan (1968) and the norms of Toglia and Battig (1978). The concreteness scale ranges from 1 to 7; with 7 being the most concrete (Roediger III, Watson et al., 2001).

Table A3

The Three 15-Word Deese-Roediger-McDermott (DRM) Lists With Backward Associative Strength (BAS) and Inter-item Associative Strength Values Used in the Present Study

King	**Inter-item**	**BAS**	**Sweet**	**Inter-item**	**BAS**	**Window**	**Inter-item**	**BAS**
queen	751	0.730	sour	434	0.405	door	191	0.156
England	20	0.000	candy	162	0.336	glass	171	0.144
crown	18	0.471	sugar	80	0.433	pane	126	0.833
prince	13	0.134	bitter	76	0.435	shade	66	0.021
George	11	0.020	good	15	0.000	ledge	11	0.152
dictator	-	0.023	taste	13	0.071	sill	49	0.682
palace	-	0.159	tooth	13	0.000	house	33	0.000
throne	9	0.759	nice	11	0.095	open	28	0.014
chess	5	0.092	honey	7	0.451	curtain	20	0.189
rule	5	0.014	soda	-	0.000	frame	12	0.014
subjects	5	0.000	chocolate	5	0.101	view	12	0.048
monarch	4	0.317	heart	3	0.000	breeze	-	0.000
royal	4	0.315	cake	2	0.027	sash	7	0.000
leader	3	0.034	tart	1	0.223	screen	3	0.027
reign	2	0.383	pie	2	0.000	shutter	-	0.480
Mean	-	**0.230**	**Mean**	-	**0.172**	**Mean**	-	**0.184**

Note. A dash represents a missing value from the Minnesota norms for (Russell & Jenkins, 1954). BAS values were based on the associative norms from Nelson, McEvoy, and Schreiber (1998). The inter-item associative strength values were based on the Minnesota norms for the Kent-Rosanoff items, which were originally derived from students' free associative responses to a presented stimulus word (the critical lure) (Russell & Jenkins, 1954).

APPENDIX B. Study Questionnaire

The following is an abridged version of the questionnaire that was given to all participants in the present study. The questionnaire included a Deese-Roediger-McDermott (DRM) task; a perception task (Ps and Qs Task); self-rated mood states; and health and demographic information.

First, thank you for participating in this study!

Please wait for the signal to begin.

Please memorize the words on the list below.

You will be asked about them later.

Do not go to the next page until you have been given the signal!

Word list

Queen

England

Crown

Prince

George

Dictator

Palace

Throne

Chess

Rule

Subjects

Monarch

Royal

Leader

Reign

=> Please wait for the signal to go on to the next page. <=

In the space below, write a short story about what is happening in this picture.

=> Please wait for the signal to go on to the next page. <=

In the space below, write down the words you memorized in the FIRST list:

Please memorize the words on the list below.

You will be asked about them later.

Do not go to the next page until you have been given the signal!

Word list

Sour

Candy

Sugar

Bitter

Good

Taste

Tooth

Nice

Honey

Soda

Chocolate

Heart

Cake

Tart

Pie

=> Please wait for the signal to go on to the next page. <=

In the space below, write a short story about what is happening in this picture.

=> **Please wait for the signal to go on to the next page.** <=

In the space below, write down the words you memorized

in the SECOND list:

Please memorize the words on the list below.

You will be asked about them later.

Do not go to the next page until you have been given the signal!

Door

Glass

Pane

Shade

Ledge

Sill

House

Open

Curtain

Frame

View

Breeze

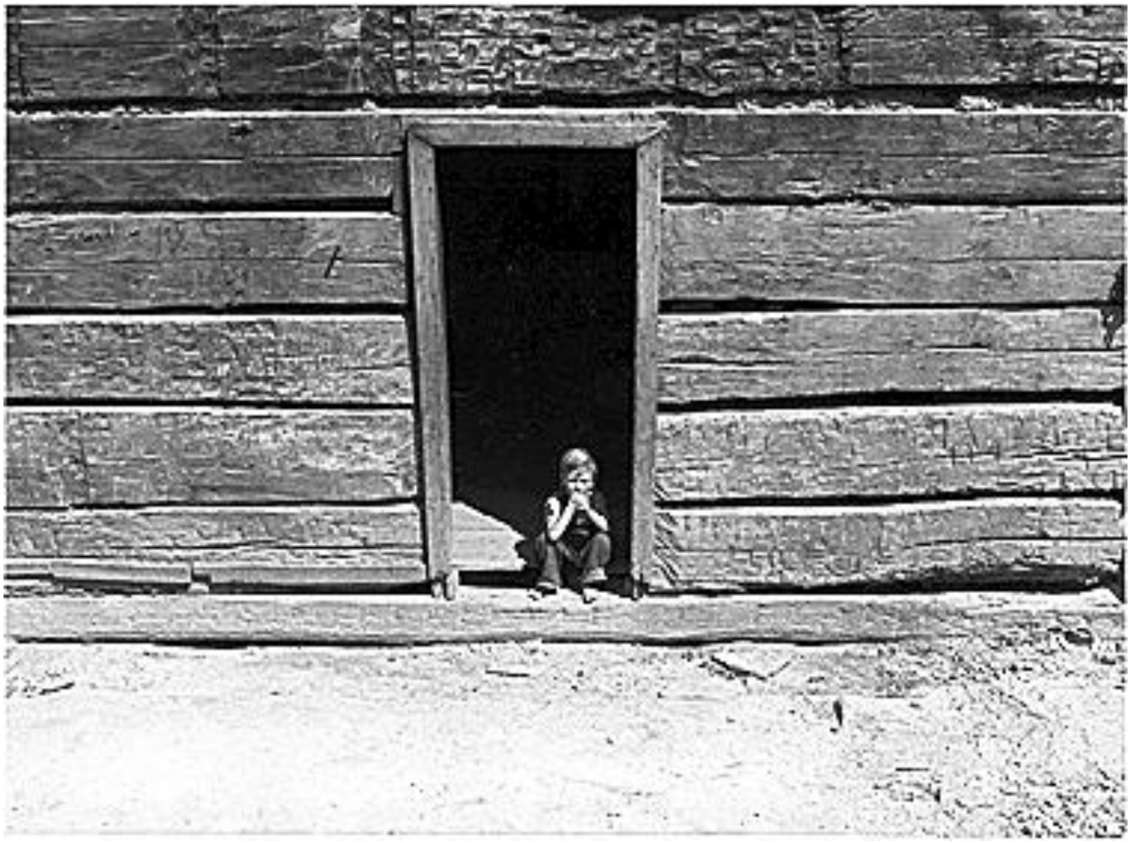
Sash

Screen

Shutter

=> Please wait for the signal to go on to the next page. <=

In the space below, tell a short story about what is happening in this picture.



=> Please wait for the signal to go on to the next page. <=

In the space below, write down the words you memorized

in the THIRD list:

Have you had experience with word lists similar to the ones used in this study?

Yes No

=> Please wait for the signal to go on to the next page. <=

Below are lines of “p”s and “q”s.

Please cross off as many qs as you can in 30 seconds.

Do not cross off any “p”s.

p p p p p q p q p q p p p p q q p q q q p p p p q q p q p q q q p p p p p q p q p p q

q q q q p p q p p p p q q p p p p q p p q q q p p p p p p p q p p p p q p q p q p q p

p p p p p p p q q q p q p q p q p p p q q q q p p p p q q q q q p p p p p p q p p p q

q q q p q q q q q p q q q q q p p p p p p q p p p p p p p p p p p p q p q p q p q p q

q q q q q q q q q q q p q q q q q q p q q p p p p p p p p p p p q p q p p p q p q p p

Next, a little bit about your day and your sleep:

Overall, what percentage of time have you been feeling in the following moods today?

In a bad mood	______%
A little low or irritable	______%
In a mildly pleasant mood	______%
In a very good mood	______%
SUM	100 %

Last night, how many hours of actual sleep did you get? ______
What time did you go to sleep? ________
Was the amount of sleep you had last night typical? Yes No

If not, do you usually get: More sleep Less sleep

How tired do you feel right now?

Very tired Somewhat tired Not tired, but not wide awake Wide awake

Do you consume caffeine (e.g. coffee, tea, soda, caffeine pills)? Yes No

If yes:

How long ago did you <u>last</u> have caffeine?

Less than 1 hour 1 hour 2 hours 3 hours 4 hours 5 hours More than 5 hours ago

When in the day do you usually consume caffeine (please circle all that apply)?

Morning Midday Afternoon Evening

What form of caffeine do you consume (please circle all that apply)?

Coffee Tea Soda Pills Other (please specify) ____________

How satisfied are you with your health these days?

Very satisfied Satisfied Not very satisfied Not at all satisfied

How stressed do you feel about *emotional* issues in your life (i.e. due to relationship problems, family problems etc.)?

Very stressed Stressed Not very stressed Not at all stressed

How stressed do you feel about *academic* issues in your life (i.e. upcoming exams, papers etc.)?

Very stressed Stressed Not very stressed Not at all stressed

How stressed do you feel about *physical* issues in your life (i.e. health, amount of exercise)?

Very stressed Stressed Not very stressed Not at all stressed

Finally, some information about who you are…

Month of birth: ___________ Year of birth: __________

Sex: Male Female

Major: ________________ College: ________________

Future occupation – please be specific: ____

Are you pre-med? Yes No

If you are pre-med, what specialty interests you the most?
____

Height (inches): _____feet_____ inches Weight (pounds):________

How physically active are you? Very Moderately Slightly Not at all

If you participate in physical activity, for which of the following reasons do you participate?
(Circle **all** that apply): Health Recreation To lose weight Other: _________________

If you play a sport, please specify: Varsity Intramural Club

How would you rate your taste perception?

Excellent Very good Good Fair Poor

How would you rate your smell perception?

Excellent Very good Good Fair Poor

What is your handedness?

Right-Handed Mostly Right-Handed Ambidextrous Mostly Left-Handed Left-Handed

Do you play a musical instrument? Yes No If yes, which instruments(s):__________________

How would you rate your proficiency at this instrument (or your best instrument if you play more than one? Expert Intermediate Beginner

Do you do puzzles? Yes No

If yes, what kinds of puzzles? ____

How often do you do them? ________________

Do you smoke? Yes No

If yes, what do you smoke (e.g. cigarettes, cigars)? __________________________

How often do you usually smoke?

Daily 2-3 times/week Once a week < Once/week

Did you smoke today? Yes No

Did you drink alcohol (including alcohol in medications) yesterday? Yes No
Did you drink alcohol (including alcohol in medications) today? Yes No
Have you ever done binge drinking (defined as drinking over an extended period of time, usually two or more days, during which you repeatedly consume alcohol to the point of intoxication, and give up your usual activities and obligations)? Yes No
Do have a cold or nasal congestion today? Yes No

Are you currently taking any prescription medications? Yes No

If yes, which ones? ____

Are you currently taking any over the counter medications (i.e. aspirin, Claritin, ibuprofen)? Yes No

If yes, which ones? ____

Do you take nutritional supplements (for example vitamin pills, calcium, protein shakes)? Yes No

If yes, which ones? ____

Do you find that these medications or supplements affect your sense of taste and/or smell? Yes No

If yes, in what way do they affect your sense of taste or smell?
____

Do you often have the feeling of having a dry mouth? Yes No

Do you have any kind of allergy? Yes No

If yes, what kind(s) do you have? ____

Do you currently have nasal polyps? Yes No Have you had them in the past? Yes No

Are you on a specific diet (i.e. vegetarian, Atkins, Kosher, South Beach, etc.)? Yes No

If yes, please specify? ____

Have you ever sustained a head injury (for example from a car accident or snowboarding)? Yes No

If yes, please specify: ____
Were you unconscious? Yes No

Circle **<u>all</u>** the groups or areas of the world from which your <u>ancestors</u> came:

Pacific Islands China Korea Japan Southeast Asia
India/Pakistan/Afghanistan Central Asia
Israel Ashkenazim Sephardim Middle East North Africa Sub-Saharan Africa
Turkey Caucasus Greece Italy Spain Portugal France Belgium Ireland
United Kingdom Netherlands Denmark Norway Sweden Iceland Finland
Latvia/Lithuania/Estonia Russia Poland Germany Austria Switzerland
Hungary Czech Republic Slovakia Bulgaria Romania
Moldavia Ukraine Pre-Columbian America

Was English the first language you learned? Yes No

If not, at what age did you start learning English? ____________

What language did you learn first? ______________

Can you read and write in this language? Yes No

For women:

What was the date of the start of your last menstrual period? _________________________

Do you use oral contraceptives? Yes No

How many periods have you had in the past 6 months? ________________________________

Are your periods regular or irregular? ____

THANK YOU SO MUCH FOR COMPLETING THIS STUDY!!!

REFERENCES

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings; and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers, 31*(4), 578-602.

Baddeley, A., Eysenck, M. A., & Anderson, M. C. (2009). *Memory*. New York, NY: Psychology Press.

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*(2), 485-502.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger III, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology, 16*(3-5), 361-384.

Balota, D. A., & Duchek, J. M. (1991). Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer type. *Brain and Language, 40*(2), 181-201.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Bauste, G., & Ferraro, F. R. (2004). Gender differences in false memory production. *Current Psychology, 23*(3), 238-244.

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Learning, Memory and Cognition, 23*(5), 1059-1082.

Bloise, S. M., & Johnson, M. K. (2007). Memory for emotional and neutral information: Gender and individual differences in emotional sensitivity. *Memory, 15*(2), 192-204.

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory.* New York, NY: Oxford University Press.

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 106*(1), 160-179.

Brainerd, C. J., Wright, 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., Reyna, V. F., & Wright, R. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110*(4), 762-784.

Buchanan, L., Brown, N. R., Cabeza, R., & Maitson, C. (1999). False memories and semantic lexicon arrangement. *Brain and Language, 68*(1-2), 172-177.
Cann, D. R., McRae, K., & Katz, A. N. (2011). False recall in the Deese-Roediger-McDermott paradigm: The roles of gist and associative strength. *The Quarterly Journal of Experimental Psychology, 64*(8), 1515-1542.
Carneiro, P., Fernandez, A., Diez, E., Garcia-Marques, L., Ramos, T., & Ferreira, M. B. (2012). "Identify-to-reject": A specific strategy to avoid false memories in the DRM paradigm. *Memory Cognition, 40,* 252-265.
Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.
Dannenbring, G. L., & Briand, K. (1982). Semantic priming and the word repetition effect in a lexical decision task. *Canadian Journal of Psychology, 36*(3), 435-444.
Deese, J. (1959a). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports, 5*(3), 305-312.
Deese, J. (1959b). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17-22.
Dewhurst, S. A., Anderson, R. J., & Knott, L. M. (2012). A gender difference in the false recall of negative words: Women DRM more than men. *Cognition and Emotion, 26*(1), 65-74.
Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks.* New York City, NY: Psychology Press.
Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory and Cognition, 38*(7), 833-848.
Gallo, D. A, & Roediger III, H. L. (2002). Variability among word lists in eliciting memory illusions: evidence for associative activation and monitoring. *Journal of Memory and Language, 47*(3), 469-497.
Hege, A. C. 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.
Herlitz, A., Nilsson, L. G., & Backman, L. (1997). Gender differences in episodic memory. *Memory and Cognition, 25(6),* 801-811.
Honda, A., & Yoshiaki, N. (2009). Sex differences in object location memory: The female advantage of immediate detection of changes. *Learning and Individual Differences, 19*, 234-237.
Howard, D. V. (1980). Category norms: A comparison of the Battig and Montague (1969) norms with the responses of adults between the ages of 20 and 80. *Journal of Gerontology, 35*(2), 225-231.

Hultsch, D. F., Masson, M. E. J., & Small, B. J. (1991). Adult age differences in direct and indirect tests of memory. *Journal of Gerontology, 46*(1), 22-30.
Jenkins, J. J., & Russell, W. A. (1960). Systematic changes in word association norms: 1910-1952. *Journal of Abnormal and Social Psychology, 60*(3), 293-304.
Johnson, M. K., Hashtroudi, S., & Stephen-Lindsay, D. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28.
Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67-85.
Kihlstrom, J. F., & Wilson, L. (1988). Reduced clustering during hypnotic amnesia for a long word list: Comment on Wilson and Kihlstrom. *Journal of Abnormal Psychology, 97*(3), 381-383.
King, J. A., Trinkler, I., Hartley, T., Vargha-Khadem, F., & Burgess, N. (2004). The hippocampal role in spatial memory and the familiarity-recollection distinction: A case study. *Neuropsychology, 18*(3), 405-417.
Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review, 1*(6), 602-609.
Kucera, H., & Frances, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.
Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology, 21*(1), 3-23.
McDermott, K. B. (1996). The persistance of false memories in list recall. *Journal of Memory and Language, 35*(2), 212-230.
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227-234.
Meyers-Levy, J., & Maheswaran, D. (1991). Exploring differences in Males' and Females' Processing Strategies. *Journal of Consumer Research, 18,* 63-70.
Murdock Jr., B. B. (1961). The retention of individual items. *Journal of Experimental Psychology, 62*(6), 618-625.
Neave, N., Hamilton, C., Hutton, L., Tildesley, N., & Pickering, A. T. (2005). Some evidence of a female advantage in object location memory using ecologically valid stimuli. *Human Nature, 16*(2), 146-163.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved from http://www.usf.edu/FreeAssociation/
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word frequent norms. *Behavior Research Methods, Instruments, and Computers, 36*(3), 402-407.
Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older

adults: Exploring the characteristics of illusory memories. *Memory and Cognition, 25*(6), 838-848.
Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement, 76* (1, Pt. 2).
Pansky, A., & Koriat, A. (2004). The basic-level convergence effect in memory distortions. *Psychological Science, 15*(1), 52-59.
Park, L., Shobe, K. K., & Kihlstrom, J. F. (2005). Associative and categorical relations in the associative memory illusion. *Psychological Science, 16*(10), 792-797.
Postma, A., Kessels, R. P. C., & Van Asselen, M. (2008). How the brain remembers and forgets where things are: The neurocognition of object-location memory. *Neuroscience and Biobehavioral Reviews, 32*, 1339-1345.
Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17*, 132-138.
Roediger III, H. L., Balota, D. A., & Watson, J. M. (2001). *Spreading activation and arousal of false memories.* Washington, DC: American Psychological Association.
Roediger III, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803-814.
Roediger III, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin and Review, 8*(3), 385-407.
Russell, W. A., & Jenkins, J. J. (1954). *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test.* Minneapolis, MN: Department of Psychology, University of Minnesota.
Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging, 8*(1), 44-55.
Seamon, J. G., Chun, L. 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 and Cognition, 30*(7), 1054-1064.
Seamon, J. G., Guerry, J. D., Marsh, G. P., & Tracy, M. C. (2002). Accurate and false recall in the Deese-Roediger and McDermott procedure: a methodological note on sex of participant. *Psychological Reports, 91*(2), 423-427.
Smeets, T., Jelicic, M., & Merckelbach, H. (2006). Stress-induced cortisol responses, sex differences, and false recollections in a DRM paradigm. *Biological Psychology, 72*(2), 164-172.

Stadler, M. A., Roediger III, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory and Cognition, 27*(3), 494-500.
Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Eribaum.
Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7*(2), 233-256.
Tse, C. S., & Neely, J. H. (2007). Semantic and repetition priming effects for DRM and unrelated study lists. *Memory & Cognition, 35*(5), 1047-1066.
Tulving, E. (1972). Episodic and semantic memory. Organization of Memory, ed. E. Tulving. W. Donaldson, pp. 381-403. New York: Academic.
Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1-25.
Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging, 13*(2), 230-241.
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.
Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review, 14*(1), 23-38.
Watson, J. M., McDermott, K. B., & Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory and Cognition, 32*(1), 135-141.