FALSE-MEMORY CONSTRUCTION: THE EFFECT OF MEMORY CONFIDENCE

Christiane N. Schneider

A Thesis Submitted to the
University of North Carolina at Wilmington in Partial Fulfillment
Of the Requirement for the Degree of
Masters of Arts

Department of Psychology

University of North Carolina at Wilmington

2004

Approved by

Advisory Committee

______________________________ ______________________________

____
Chair

Accepted by

____
Dean, Graduate School

TABLE OF CONTENTS

ABSTRACT

According to Roediger and McDermott (1995), a false memory is a memory of an event that never occurred. A large body of research has explored the false memory effect and the factors that influence false memory production. The purpose of the present study was twofold. The first aim was to examine the extent to which a participant's confidence in their overall memory ability influences the production of false memories. The second aim was to explore the extent to which fluctuations in memory confidence potentially influence fluctuations in the false memory phenomenon. To these aims, participants were randomly assigned to one of three groups: high-confidence group, low-confidence group, or neutral group. Participants received positive, negative, or neutral feedback about their memory performance on three different memory tests in an attempt to experimentally manipulate participants' confidence in their memory. Using the Roediger and McDermott (1995) paradigm, participants were administered a total of 32 word lists and were administered a recall test after each list was presented. Each list contained 15 words associated with one non-presented word (critical lure). After all 32 lists were presented, participants completed a recognition test in which they were asked to identify the words presented on each list and to make remember, know, and guess judgments (Tulving, 1985). The analysis on the recall and the recognition test revealed a false memory effect: studied items were recalled and recognized at a higher rate than critical lures which in turn were recalled and recognized at a higher rate than non-critical intrusions or new words. No significant differences between the three memory manipulation conditions were observed, indicating that the memory manipulation did not affect false memory production.

ACKNOWLEDGMENTS

I would like to thank my parents, for without you, I undoubtedly would not have made it this far. Your love and support from a continent away has given me the strength to push on when I didn't think I could make it. It has been a tough road, but with you two behind me, I've made it to the finish line.

I would also express my appreciation to my committee members Dr. Lisa Jenkins, Dr. Nora Noel, Dr. William Overman and Dr. Ray Pitts. Thank you all for your patience and understanding, as well as all the help you have given me to see this thesis to its completion.

Finally, I need to thank my wonderful friends who have been incredibly supportive. Thank you, Tafline, for listening to me for hours and guiding me when I most needed you. Thank you, Stacy, for putting up with me and being an incredible roommate. And to everyone else in Wilmington, I will never forget you!

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Research in memory has been guided by Bartlett's (1932) assertion that memory is reconstructive rather than photographic in nature. In one experiment, Bartlett (1932) had participants listen to a North American folk-tale called "A War of the Ghosts." After a period of time, participants were asked to recall the story as accurately as possible. Bartlett found that participant's recollections typically contained the key information of the folk-tale, but the specific details of the story were often either missing or incorrect. After further research, he proposed that humans do not have photographic memories and that information retrieved from memory is not always accurate.

On the basis of his research, Bartlett (1932) argued that memorial information is reconstructed using information from both the actual event and an individual's previous experience with similar events. The exact (or verbatim) information is used to form a general framework in which the specific details are incorporated. Bartlett argued that, because the specific details of an event are likely to be forgotten quickly, humans probably generate the details of an event using memories of similar experiences. Thus, when an individual is attempting to recall or recollect an event, the framework is rarely missing or erroneous, but the specific details may be forgotten or incorrect.

The reconstructive nature of memory has also raised issues outside the laboratory. Psychologists and psychiatrists have used hypnosis to help their patients or clients "remember" traumatic events that happened to them during their childhood. This use of hypnosis has become a hot topic, as some researchers have suggest that some of "recovered" memories by patients undergoing therapy are the result of suggestive or leading statements made by the therapist while the patient underwent hypnosis (e.g., Steblay, Mehrkens & Bothwell, 1994). In some unfortunate circumstances, patients have falsely accused family members, friends, or acquaintances of crimes that were never committed (e.g. sexual abuse). In recent years, the accuracy and appropriate use of eyewitness testimony has also been called into question. The current consensus is that there are many factors that influence the accuracy of an eyewitness's testimony (Cutler, Penrod and Martens, 1987b; Kassin, Ellsworth, & Smith, 1989; Shapiro & Penrod,1986), and that at least under certain circumstances, eyewitness accounts can be wholly inaccurate (Migueles & Garcia-Bajos, 1999).

James Deese (Deese, 1957; Deese, 1959; Deese & Hardman, 1954) was the first to conduct laboratory investigations of false-memory construction. In his research, participants were administered 36 word lists that contained 12 items each. Each list contained 12 words all semantically related to one non-presented word. For example, one list included the words hard, light, pillow, plush, loud and cotton and the highly associated word that was not presented on the list (i.e., the critical lure) was soft. The lists were randomized in five different orders, and presented auditorily. Participants orally recalled the words that were presented on each list. Deese calculated the percentage the critical lure was recalled and found that some lists produced few critical lures (butterfly = 0%) while other lists produced the critical lure at a much higher rate (sleep = 44%). He hypothesized that participants recalled the critical lure because the words on the lists activated other highly associated words not presented on the lists. The participants then recalled these highly semantically related words thinking that they actually appeared on the lists.

In order to explore this hypothesis, he administered a free-association test consisting of all the words from the 36 lists. He instructed participants to report the first word that came to their mind for each word that was presented from the 36 memory lists. He calculated the association strength for each word by summing the number of times the critical lure was paired with each word. The mean association strength for the entire list was calculated by summing the strength of the each word's association to the critical lure and dividing that sum by the number of words on the list. Using this information, Deese calculated the relationship between the frequency that the critical lure was recalled to the mean association strength of the lists to establish the probability that critical lure would be recalled for each list. He found a positive linear correlation between the mean association strength of a list, and the recall of the critical lure (P Recall of Critical Lure = 3.7 + 1.63*mean association strength of a list; $r = .873$). Using this relation, if the mean association strength of a list was 60%, the probability of the critical lure being recalled would be 1.0. On the basis of the strength of the relation between the recall of critical lures and the mean association strength, Deese concluded that the words on each memory list activated other semantically related words that were subsequently incorrectly reported as studied words at recall (Deese, 1959).

Recently, Roediger and McDermott (1995) sought to replicate Deese's (1959) findings using 6 lists that produced the largest number of critical lures in Deese's study in an attempt to extend their findings to recognition memory. In the first experiment, participants were read words one a time from a word list. At the end of each list, participants were told to write down as many words from the list that they could remember. After all the lists were administered, participants completed a recognition test that contained two weakly related words from each of the studied lists, the critical lures from each list, and several unstudied words. For each word, participants were also instructed to rate that word from 1 (sure it was new) to 4 (sure it was old) to indicate whether each word presented was on one of the studied word lists or if it was a new word.

Results revealed that the studied words were recalled at a rate of .65, and the critical lures were recalled at a rate of .40 (Roediger & McDermott, 1995, Experiment 1). Words that had not been presented on the lists were recalled at a rate of .14. These results indicate that the studied words were recalled more frequently than the critical lures, and that the critical lures were produced at a significantly higher rate than the non-studied words. Similar results were obtained on the recognition test. That is, studied words rated as either a 3 or 4 (old) were recognized .86 of the time, critical lures rated as either a 3 or 4 were recognized .84 of the time, and words that had not been presented (new words) were recognized as old .02 or the time. Although critical lures were not recognized at the same rate as studied items, critical lures were recognized as being old significantly more often than items that had not been presented on the lists. Taken together, these findings replicate Deese's (1959) results and indicate that the false memory phenomenon can be observed in both recall and recognition.

In a second experiment, Roediger and McDermott were interested in determining whether participants identified critical lures as 'old' on a recognition test because they actually remembered hearing the critical lures or because they simply were responding to a feeling of knowing that the critical lures were on the lists. In their experiment, they used word lists similar to those constructed by Deese (1959), except that each of the 15 words in each list was highly semantically related to one word that was not presented on the list. For example, one list included the words hard, light, pillow, plush, loud and cotton and the highly associated word that was not presented on the list (i.e., the critical lure) was soft.

After each list was presented, participants either performed a recall test or completed a math test. The inclusion of a math test after half the lists enabled Roediger and McDermott to determine whether recalling a list immediately after it had been presented would affect recognition performance. After all 16 lists were presented, participants completed a recognition test similar to that presented in their original experiment and indicated whether each word was “old” or “new”. In addition, for all items rated as 'old', participants indicated whether they actually remembered the word being presented (remember) or if they just knew the word was in the list (know). The remember/know procedure, introduced by Tulving (1985), was adapted to examine the nature of the memory for the critical lure in order to explicitly determine whether the information participants recalled or recognized was retrieved from specific representations in memory or simply from strong feelings of familiarity.

Results revealed that studied words were recalled at a probability of .62 while the critical lure was recalled at a probability of .55, replicating the results obtained in Roediger and McDermott's (1995) first experiment. For those participants who recalled the studied items immediately following presentation, studied items were correctly recognized at a rate of .79 and these were identified as remembered at a rate of .57. The critical lures were correctly recognized at a rate of .81. For those participants who performed the arithmetic test immediately after each list was presented, participants correctly recognized the studied items at a rate of .65 and identified them as remembered at a rate of .41. The critical lures from the lists that were followed by arithmetic problems were incorrectly recognized at a rate of .72 and identified as old at a rate of .38, and identified as being remembered at a rate of .38. Not surprisingly, results revealed that recalling information immediately after it is presented facilitates later recognition of that information as well as increasing the probability of remembering that word as having been presented on one of the word lists.

On the basis of the results from both experiments, Roediger and McDermott (1995) concluded that the pattern of memory performance that they obtained could be explained using a more general theory of memory called spreading activation. In this theory, memory is represented as a framework of interconnected nodes. A node is essentially a concept or a piece of information that is linked to several other concepts. Consider the word soft. This word is highly interrelated to other words such as hard, cotton, and loud. As such, each of these words would be represented as a separate node in a network and the relationships between these words would be reflected by their connections. The links between nodes are created on the basis of experience with those words and the strength of the connection is based on various criteria that designate the importance of that relationship. For the present discussion, it is sufficient to assume that each of the connections between the words is of the same strength. One possible network that specifies these relationships is portrayed in Figure 1. Hard is connected with soft and cotton because they are semantically associated with one another (because connections in a network of this type can represent synonym or antonym associations). However, because loud is only associated with soft, there are no connections between loud and any of the other words in the network.

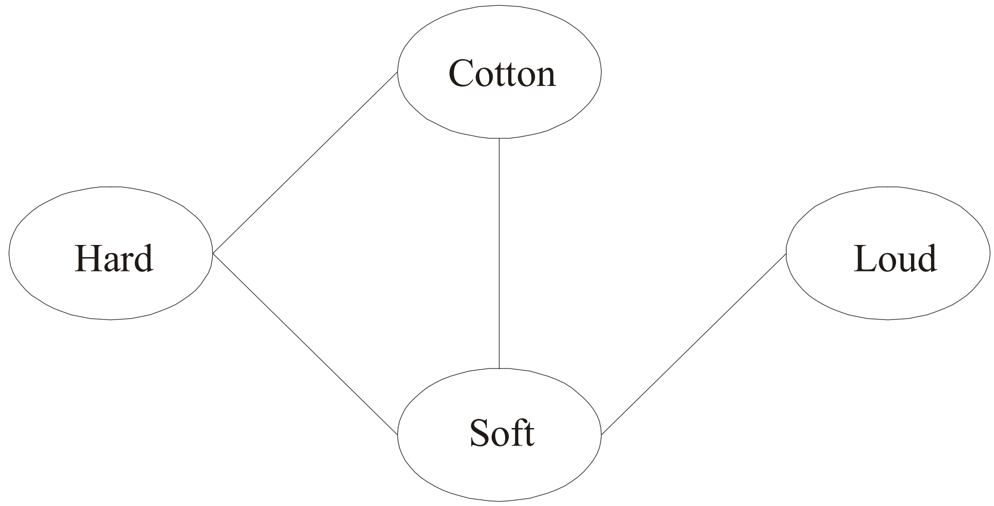


Figure 1. Example of a spreading activation network for the word ‘soft’.

The process of memory retrieval begins with an input, which activates one node. Activation then spreads from that node to other nodes that are linked to it. Nodes that are not linked to the nodes that become activated are unaffected. As other inputs activate other nodes in the network, activation spreads from those nodes to the other nodes that are connected to them. Each activation of a node leads that node to become more and more activated such that all the activations sum together. The most highly activated node(s) in a network will comprise what is recalled or recognized at any given time. If too much time passes between activation of information in a network, retrieval or recognition may not be possible because activation in a network decays exponentially (Bechtel & Abrahamson, 1991).

Roediger and McDermott (1995) used the theory of semantic activation to explain their results. Specifically, they argued that each word in a list and any high-associates of those words (e.g., soft) would be nodes in an interconnected network. As each word in the list is activated, activation would also spread to other connected nodes, including the critical lure node, 'soft'. Because all the words in the word lists administered were high associates of the word soft, that node would receive large amounts of activation over the course of a list. After all the words in a list are presented, the node for each presented word and the node for the critical lure would have some level of activation above zero. At recall or recognition, any node that is activated above some threshold amount will be recalled or recognized as having been on the list.

Although Roediger and McDermott (1995) used the spreading activation theory to explain their data, several other theories have been proposed to explain why false memories occur. Specifically, the implicit activation hypothesis (Underwood, 1965), the source monitoring theory (Johnson, Hashtroudi & Lindsay, 1993), the fuzzy-trace theory (Brainerd & Reyna, 1990), and the criterion-shift theory (Miller & Wolford, 1999).

Implicit Activation Hypothesis

Underwood (1965) proposed the implicit activation hypothesis to account for false-memory construction. Similar to the spreading activation theory, Underwood's hypothesis concentrates on the nature of the processing of semantically associated words that are not presented on the word lists. Specifically, Underwood argues that the critical lure is unintentionally or unconsciously activated as a result of activating words that are highly associated with it, whereas Roediger and McDermott (1995) do not make any assumptions regarding the conscious or unconscious activation of this information. According to Underwood, the critical lure and other non-presented words are encoded during the list presentation as if they had been presented. Thus, in recall and recognition tests, the critical lures should act similarly to studied items. For example, when an individual sees a list that contains the words, hard, light, pillow, plush, loud and cotton, the word soft is automatically, but implicitly activated. At recall, the individual should write the word soft down because it was encoded in the same fashion as all the other words in the list. At recognition, the individual would recognize the word soft as being a studied item for the same reasons.

The implicit activation theory would uniquely predict that false memories would still be generated in experiments in which the word lists are presented below participant's conscious level of encoding (e.g., 50 ms.). This prediction has been tested recently by Seamon, Luo, and Gallo (1998). They tested Underwood's hypothesis that critical lures are unconsciously activated using two methods that would minimize the explicit activation of the critical lure: near-threshold presentation rates and concurrent memory load. In the first experiment, two groups were tested: the concurrent memory load group, in which participants were given a digit sequence to memorize before the first list was given; and a control group, in which participants received no digit sequences to memorize. Participants in both groups saw words presented for 20 ms, 250 ms, and 2 s. After all the lists had been presented, participants were administered a recognition test.

Results revealed that hit rates were lower for those in the concurrent memory group than those in the control group. Correct recognition was also lower at near threshold presentation rates (20 ms) than slow presentations (250 ms and 2 ms). With regard to false recognition of the critical lure, the concurrent memory group showed lower false recognition rates than in the control group when the presentation rate was 250 ms and 20 ms. This is an important finding because it suggests that the false memory phenomenon is, at least to some extent, the result of an implicit activation of critical lures.

In a second experiment, Seamon et al. (1998) attempted to replicate these findings using a within-subjects design. In this experiment, participants were administered four lists under concurrent memory load conditions and four lists under normal conditions. Not surprisingly, hit rates were lower for lists under concurrent memory loads than lists under normal conditions. Correct recognition for studied items and false alarms for critical lures was lower at shorter presentation rates than at longer presentation rates, consistent with the findings from Experiment 1. However, critical lures were still falsely recognized at all presentation rates. On the basis of this research, Seamon et al. (1998) concluded that these results could only be obtained if the words on the lists activated the critical lure unconsciously, as Underwood (1965) suggested.   In other words, the activation of the both the words and the critical lure were done automatically, without any conscious thought from the participants.

As the participants would not be able to consciously encode the words from the list, no differences between remember and know judgments would be predicted under the implicit activation hypothesis framework.

Source Monitoring Theory

The source monitoring theory (Johnson et al., 1993) is a framework that is based on a more general theoretical framework of memory functioning. From this perspective, each event is stored in memory in conjunction with certain perceptual (e.g., color, texture, or sound), contextual (e.g., spatial or temporal), and reflective or affective (e.g., emotional reactions or semantic associations) characteristics. For example, the experience of hearing the word ‘hard’ might be stored in memory with other characteristics such as it’s position in the list, the tone of voice of the experimenter, and any images that may have been spontaneously produced when the participant heard the word.

Johnson et al., (1993) argued that the ability for a participant to determine whether a specific word was presented on a list involves the activation of that memory and the activation of any additional characteristics stored with that memory. When probed about an event, a little or a lot of additional information may be necessary in order to judge whether that event was real, imagined, in the recent past, or in the distant past. The more characteristics a memory has associated with it, the greater the likelihood that it will pass the source monitoring test and be considered a memory of a true experience.

Johnson et al. (1993) are careful to point out that the quality of the information stored in memory at encoding is critical to later retrieval. If the information at encoding is inaccurate or invalid in some way, the source monitoring process may go awry and yield faulty conclusions about later retrieved information.

According to the source monitoring theory, false memories can result from two underlying factors (Johnson et al., 1993). Participants may confuse the additional characteristics of true memories and may generalize them to related information that was not presented. Alternatively, there may be individual and/or developmental differences in the extent to which participants can store and effectively use such additional characteristics in order to accurately judge the source of a memory. In either case, the source monitoring theory would posit that the false memory phenomenon is a result of the faulty or inappropriate use of the additional characteristics of information stored in memory. In terms of remember and know judgments, participants would report that they remember an item in situations where source information (correct or incorrect) is available. On the other hand, know judgments would be made even when the source information is weak.

Fuzzy Trace Theory

Brainerd and Reyna (1990) have argued that memorial information is encoded and stored in two forms: general gist and verbatim. General gist memory consists of information about the overall meaning of an event whereas verbatim memory consists of the exact details of an event (Brainerd et al, 1990). Consider a situation in which participants study a list of words that are all semantically related. In this situation, individuals form a gist representation based on the semantic content of those words while simultaneously encoding the exact details of the information being presented.

On recall or recognition tests, individuals can identify words that are remembered verbatim (e.g., light, pillow, cotton), words that semantically fit the gist (e.g., soft, feather, white), or both. From this perspective, Brainerd and Reyna (1990) have argued that false memories or memories for information not presented result from an individual relying on gist-based information rather than verbatim information in memory. In terms of remember and know judgments, the participant would indicate they remember the item when it is based upon verbatim traces, and alternatively, would indicate they know the item was present when the decision is based more on gist traces.

Criterion Shift Theory

Criterion shift theory is concerned with variations in the decision processes that are used when individuals make recall or recognition decisions about information stored in memory. In essence, this theory proposes that all participants have an implicit set of rules or guidelines that they follow when retrieving or recalling information from memory. Such criteria can be altered by the individual, through experimenter instructions, or by the situation at hand.

For example, suppose a doctor is asked to recall what medication should be administered to a patient who has diabetes. In this circumstance, because the consequence of making a mistake in recollection is high, that doctor would likely set a very high criterion on the retrieval of all possible drug information in order to ensure that the correct drug is retrieved. In contrast, suppose a 90-year-old adult is asked whether he/she had chocolate cake for his/her 5th birthday party. The consequence for recalling incorrect information is small, and therefore, a less stringent criterion would be set so that any potential information about that event would have a chance to be recalled.

When participants are asked to recall or recognize information presented in a word list, they will adopt some criterion between one that is strict and one that is lenient. Participants who adopt a stricter criterion will recall only those words that they are absolutely sure that they remember hearing whereas participants who adopt a more lenient criterion may recall words from the list in addition to other words that were not presented. According to Miller and Wolford (1999), false memories in the Roediger and McDermott (1995) paradigm result from participants using a less stringent criterion for recall and recognition as a result of the nature of the word lists. That is, because the word lists contain items that are highly semantically related to one another, participants may be more likely to say that the critical lure is old simply because it is related to other words that were on the list. Therefore false alarms to critical lures represent individual differences in the criterion used for recall and recognition of memorial information and not the production of a false memory. Within an individual, the number of remember and know judgments would fluctuate depending on how liberal a criterion that individual adopts. Setting a strict criterion, would yield more remember responses whereas setting a lenient criterion would results in more know judgments.

Factors that Influence False Memory

A considerable amount of research has been conducted to determine the factors that influence false-memory construction. A discussion of the major factors is presented below.

Presentation

As mentioned above, Seamon et al., (1998) tested Underwood's (1965) hypothesis that critical lures are unconsciously activated using two methods that would minimize the explicit activation of the critical lure: near-threshold presentation rates and concurrent memory load. Their two experiments showed that critical lures were still recalled when presentation rates did not allow for correct recognition of studied items. This lends support for the idea that critical lures are, to some extent, due to implicit activation.

Modality

Several studies have examined how the manner in which the lists are presented affects false-memory construction. Smith and Hunt (1998; Experiment 1) used the 5 most highly associated words from 10 randomly chosen lists from Roediger and McDermott (1995). The words were presented to participants as one 50-word list. The list was presented to the participant either aurally or visually, which was then followed by a written free recall test. The results indicated that when the words were presented aurally, recall of the studied words and critical lures did not differ significantly. On the other hand, when the words were presented visually, studied items were recalled at almost twice the rate as critical lures. In a second experiment, Smith and Hunt (1998; Experiment 2) sought to extend these findings to a recognition test. Twelve words from the same six lists were presented either visually or aurally, and after each list, participants recalled as many words as they could. A recognition test was administered at the conclusion of the list presentation. The results of the recall and recognition test replicated the results from Experiment 1. Specifically, studied items were recognized at a higher rate than critical lures when the words was presented visually, but recognition did not differ between studied items and critical lures when the words were presented aurally. These results were the first indication that the modality of the presentation of the lists had an effect on false-memory construction.

Kellog (2001) further investigated the results from Smith and Hunt (1998) by manipulating the test modality of the paradigm as well as the presentation. Sixteen lists from Roediger and McDermott (1995) were split into four groups of four lists and were either presented aurally or visually in a within group design. After the visual or aural presentation of each list, participants were asked to recall as many words as possible by either writing the words down or saying them out loud. The results replicated the findings from Smith and Hunt (1998) for written recall: fewer critical lures were recalled when the lists were presented visually than when they were presented aurally. However, no difference was observed in the rate of recall of critical lures when recall was spoken rather than written. Thus, the presentation modality did have an effect on the recall of critical lures, whereas the recall modality did not.

In a similar study, Maylor and Mo (1999) used 12 lists from Roediger and McDermott (1995) and created a recognition test for each of the 12 lists. Lists were either presented visually or aurally and the recognition test was administered either visually or aurally in a with-in subjects design. The results showed that correct recognition of studied items was higher with auditory presentation than with visual presentation. On the other hand, false recognition of the critical lure was higher with visual presentation than with auditory presentation. Furthermore, false recognition rates were highest when the study and test modalities differed, for example, if the words were presented visually and the recognition test was auditory.

Based upon the source monitoring theory, participants should be most accurate in recalling or recognizing studied words when the recall or recognition test is administered using the same modality as the modality in which the words were presented because the source is the same? (Tulving & Thompson, 1973). False recall and recognition of critical lures should be highest when the modality of the presentation and the test are opposite, as was shown by Maylor and Mo (1999).

Gallo, McDermott, Percer and Roediger (2001) attempted to shed light on these conflicting findings of presentation modality. For all three experiments in their study, they followed the Roediger and McDermott (1995) paradigm, yet manipulated the presentation and test modality as well as the between/within subjects design. In Experiment 1, 24 lists were presented to the participants, and half were followed by math problems and half were followed by immediate recall. At the conclusion of the recall portion, participants were given the recognition test which contained words from positions 3, 8, 10, and the critical lure from the lists that were presented as well unstudied items. They were asked to indicate if the word had been presented and to make the remember/know judgment. They were then given the same recognition test, yet this time they were asked to indicate if the word had been presented or not, and if it had been presented, if it was presented auditorily, visually, or if they did not know.

The results showed that the study modality had an effect upon critical lures, as they were recalled at a higher rate following auditory presentation than after visual presentation, replicated Smith and Hunt's (1998) findings. This was also found with the recognition of critical lures, although it was smaller than the effect found in recall. Analysis on the remember/know judgments revealed that studied items and critical lures were judged as remembered more often when the corresponding list had been presented auditorily than visually. Results from the modality judgment revealed that participants were quite accurate in their judgment. In addition, critical lures were categorized as having been presented in the same modality as the corresponding list, and this was done at the same level of studied items. The results replicated those from Smith and Hunt (1998) although with a smaller effect.

Experiment 2 was designed to examine if test modality had an effect on the recall and recognition of studied items and critical lures. Twenty-four lists were presented to the participant, either visually or aurally in a blocked design, so that 12 lists were presented aurally followed by 12 presented visually, or 12 lists were presented visually followed by 12 presented aurally. After the presentation of the lists was complete, the participant was given the recognition test. The recognition test comprised of words 1, 8, 10, and the critical lure from lists presented in addition to new items. Half of the items from the lists that were presented aurally were presented visually on the recognition test and the other half were presented aurally on the recognition test. In addition, half of the items from the lists that were presented visually were presented aurally on the recognition test and the other half were presented visually. Half of the new items were presented aurally and the other half visually. The remember/know judgments were expanded, and participants were asked to indicate if the remembered the item, knew the item had been presented, or if they were guessing the item had been presented, and they had no feelings of remembering or knowing.

The results indicated that test modality did have an effect on studied items and critical lures. False recognition of critical lures was higher with visual presentation than with auditory presentation. In addition, critical lures visually presented were falsely recognized at a higher rate with visual test than when the critical lures were presented visually and then auditorily tested. When the critical lures were presented auditorily, there was no difference in the test modality. Correct recognition of studied items was higher with visual presentation than with auditory presentation. In addition, studied items visually presented were correctly recognized at a higher rate with visual test than when the studied items were presented visually and then auditorily tested. When the studied items were presented auditorily, there was no difference in the test modality. Further, the rate of correct recognition of studied items was higher than the false recognition of critical lures in every case, except when the list was presented auditorily with visual test. Analysis of the Remember/Know/Guess judgments revealed that Remember judgments of studied items were greater than Remember judgments of critical lures except when the lists were presented aurally with a visual recognition test, when they were judged as remembered at the same rate. Again, these results are consistent with the findings of Smith and Hunt (1998) and are not consistent with the findings of Maylor and Mo (1999).

Experiment 3 was an attempt to replicate the findings from Maylor and Mo (1999). The procedure was the same as in Experiment 2, yet with a between-subjects design. Participants were randomly assigned to one of four conditions: auditory presentation-auditory test, auditory presentation-visual test, visual presentation-visual test, visual presentation-auditory test. The procedure and recognition test was the same as in Experiment 2.

As in Experiment 2, critical lures were falsely recognized at higher rates in visual presentation than with auditory presentation, and studied items were correctly recognized with visual presentation than with auditory presentation. There were no statistically significant findings concerning critical lures between the four conditions, although this was most likely due to the weaker between-subjects design. However, for studied items, when the presentation modality and the test modality were the same, correct recognition of studied items was higher than when they were not the same. This result was partially consistent with Experiment 2, as visual presentation with visual test produced higher rates of correct recognition than all the other conditions. Consistent with Experiment 2, the rate of correct recognition of studied items was higher than the false recognition of critical lures in every case, except when the list was presented auditorily with visual test. Analysis of the Remember/Know/Guess showed that Remember judgments of studied items were greater than Remember judgments of critical lures except when the lists were presented aurally with a visual recognition test, when they were judged as remembered at the same rate.

The spreading activation theory, fuzzy-trace theory, and the implicit activation theory would have difficulty explaining these results, as the presentation modality should have little effect upon the level of activation in both of these theories. On the other hand, Gallo et al. (2001) have argued that the source monitoring theory can explain the consistent findings of Smith and Hunt (1998), Kellog (2001), and Gallo, McDermott et al. (2001). Specifically, it is posited that participants have better source monitoring ability with visual presentation than auditory presentation because items presented with a visual modality are better encoded and more salient than in the auditory modality.

However, the variability of the findings warrants more research before a good explanation would fit any of the theories.

Distinctiveness

The distinctiveness heuristic is based upon the source monitoring theory in which individuals use decision strategies when evaluating information by trying to attribute the information to a source. It has been shown that individuals shown pictorial stimuli make their decisions on the basis of their ability to attribute it to a source (reference). Dobson and Schachter (2001) wanted to extend these findings to the Roediger and McDermott (1995) paradigm. In Experiment 1, they had participants view 18 lists from Roediger and McDermott (1995) and half the participants were instructed to say the word on the screen and the other half heard the word being said while it was on the screen. All participants were then administered a visual recognition test

The results showed that critical lures were falsely recognized at a higher rate when the participants heard the words than when they said the words, although there was no difference between the two conditions in the rate of recognition of studied items. In the hearing condition, studied items and critical lures were recognized at the same rate, although in the saying condition, critical lures were recognized at a significantly lower rate than studied items. Yet, participants in the say condition did not show a reduction of false recognition of new items, which would have been expected by distinctiveness heuristic, leading Dobson and Schachter (2001) to believe that the distinctiveness heuristic was being used in some situations and not others.

To test this idea, Dobson and Schachter (2001) created a within subjects design in Experiment 2, in which 8 lists were heard by the participant and 8 lists were spoken. The results from this experiment showed that participants in the say condition had a lower rate of false recognition of new items and higher rates of correct recognition of studied items compared to participants in the hear condition. Yet the two conditions did not differ in the rate of false recognition of critical lures. Their results indicated that participants had difficulty using the distinctiveness heuristic when two types of encoding were used. These results, taken together with the results from Experiment 1 indicate that the distinctiveness heuristic is one decision strategy that can be used by participant in the Roediger and McDermott (1995) paradigm.

As Dobson and Schachter (2001) were studying an aspect directly related to the source monitoring theory, their results can be explained, albeit not as neatly, by other theories. The spreading activation theory, the fuzzy-trace theory, and the criterion shift theory would all presume that different strategies, or thresholds, could be utilized when making decisions in the Roediger and McDermott (1995) paradigm. Although all three theories would be able to explain the results from Experiment 1, none of them could adequately be able to predict the results from Experiment 2.

Gallo, McDermott, Percer and Roediger (2001) were also interested if participants used a distinctiveness heuristic in Experiment 3 (mentioned above). They argued that when modality effects are manipulated in a within subjects design, the participant relies on list-specific information, while a between-subjects design would allow the participant to make decisions based upon both list-specific information and the distinctiveness heuristic, thereby reducing false recognition of critical lures. The distinctiveness heuristic is able to somewhat explain the results from the three experiments. In the case of visual presentation, participants are less likely to falsely recall or recognize the critical lure because there was no visual recollection for them. Yet there were no differences in the rates of false recognition of the critical lure between Experiment 2 (within subjects design) and Experiment 3 (between subjects design), a finding that is contrary to the predictions made with the distinctiveness heuristic. This led the authors to conclude that participants based their decisions on list-specific modality information rather than the distinctiveness heuristic.

Delay

Previous studies have investigated the effects of a delay before a free recall test, yet little was known about how a delay would affect false-memory construction in the Roediger and McDermott (1995) paradigm. McDermott (1996) designed an experiment in which a delay was imposed before the administration of a recall test. Specifically, immediately after being presented with word lists, one group of participants recalled as many of the words in the list as they could in 90 s and then completed math problems for 30 s. The other group of participants completed 30 s of math problems followed by a 90-second recall test. An additional free recall test was administered 48 hours after the first testing session was completed.

Results revealed that the 30-second delay before recall lowered the correct recall of studied items, but it did not affect the false recall of critical lures, consistent with results obtained by Roediger and McDermott (1995). Results from the free recall test two days later revealed that critical lures were recalled at a higher rate than studied items. This result is interesting because one might predict that recall at long delays would be poorer than at short delays due to simple decay of trace information in memory. No theory of decay can explain why critical lures are retrieved at a higher rate than studied items.

Thapar and McDermott (2001) further examined the effects of delay on false memory construction by also manipulating the level the participants processed the words on the list. In Experiment 1, participants were directed to rate the pleasantness of the meaning (deep processing) of the word for 8 lists, the number of vowels (shallow processing) in each word for 8 lists, and the color (shallow processing) of the words for 8 lists as they were presented. Participants were then asked to immediately do a free recall test, or asked to return in either 2 or 7 days to complete the free recall test. Results from Experiment 1 indicated a significant decrease in the rate of recall of studied items and critical lures over time, although this was more dramatic with studied items than with critical lures. Initially the probability of recalling a studied item was higher than recalling a critical lure (immediate) although after 7 days, critical lures were recalled at higher rates than studied items. In addition, lists which were deeply processed were more likely to produce the studied items and critical lures at all three time intervals than lists shallowly processed.

A second experiment was conducted to replicate and extend these findings using a recognition test. The procedure was similar and the results showed the same trends observed in the recall data in the recognition data. Thus, Thapar and McDermott (2001) revealed that recall and recognition of studied items and critical lures decreased over time, yet more so for studied items than for critical lures. In addition, the level of processing had a profound effect on recall and recognition at all three time intervals.

Fuzzy trace theory would predict these results because although the details have been lost due to decay, the gist remained strong enough for recall and recognition. At test, individuals could then rely more on the fuzzy trace than the details or verbatim memory. The source monitoring theory would also predict this result if one assumes that participants lower their threshold on which they base their memory judgments. The longer the delay from list presentation to test, the greater the amount of decay. To outweigh the effects of decay, participants make adjustments for the change in activation and adopt a lower threshold in order to remember some information about the event. This would account for the remembering critical lures, as less emphasis is placed on remembering exactly if a word was presented, and more on emotional and cognitive features of a word – features the individual associates with the critical lure. Similarly, the criterion shift theory would predict McDermott's (1996) results if one assumes that participants lower their criterion in order to remember some information about an event that occurred several days earlier. However, the spreading activation theory has difficulty predicting this result because both the words on the lists and the critical lures were highly activated during encoding, and there is no a priori reason to expect that one should decay more rapidly than the other.

Repetition

Recently, McDermott (1996, Experiment 2) examined the extent to which multiple presentations of the same lists would have an effect on the number of studied items and critical lures recalled. Using 3 lists from Roediger and McDermott (1995), two lists of 45 words were generated that presented the 45 words in two different sequences. In one list, the words from the 3 chosen lists were presented one after the other in the order in which Roediger and McDermott (1995) used (blocked order). In the other list, the words from the chosen 3 lists were presented in a completely random order (random order). Participants saw either the blocked or random list five times and were asked to recall the words from the list after each presentation. Participants returned the next day to complete a free recall test.

Results revealed no effect of blocked or random order on recall of studied items or critical lures. However, recall for studied items increased while the number of critical lures reported decreased as list presentation increased. Interestingly, the false memory effect was still robust even after 5 list presentations. After a 24-hour delay, recall for studied items decreased while recall of critical lures increased. These results indicate that although repetition reduces the magnitude of the false memory effect, it does not eliminate the phenomenon. Such a finding may be best explained by the fuzzy trace theory if one assumes that participants learn to rely on verbatim traces with increasing list presentations. That is, as participants see the same list multiple times, they may learn to adopt a verbatim recall strategy and abandon the influence of a gist-based strategy. Because false recall of the critical lure is hypothesized to result from the reliance on gist-based trace, one would expect the pattern obtained by McDermott (1996): higher rates of correct recall of studied words and lower rates of false recall as list repetition increases. It is unclear how any of the other theories of false-memory construction (at least as they are currently formulated) can account for this result.

Explanation

Research on perceptual illusions (e.g., the Müller-Lyer illusion) has shown that conscious knowledge of the illusion does not diminish the effect (Gregory, 1968).

Several researchers (e.g., Gallo et al., 1997; McDermott & Roediger, 1998) have been interested in determining if the same phenomenon would hold for false memories. McDermott and Roediger (1998) performed three different experiments using the Roediger and McDermott (1995) paradigm in which participants were informed of the false memory effect. Specifically, participants were told that they would hear lists of words to remember, each list containing words associated to one specific word. Some lists would contain that word while others would not, and participants were told that they must first figure out the word and then decide if it was presented or not. Participants were then given the word list for 'king' as an example. Results showed that warning participants about the false memory effect lead to a reduction, but not an elimination of the false recognition of information.

Gallo et al (1997) used a more direct method to study the extent to which warning participants about the phenomenon would lead to the elimination of false memories. Sixteen lists from the Roediger and McDermott (1995) paradigm were used. Participants were placed in one of three different experimental groups: an uninformed group, in which participants were told to try and remember as many words as possible because they would be given a recognition test later; a cautious group, in which participants were told that some words on recognition test would be new but would be very similar to some words that had been presented in the lists; and a forewarned group, in which participants were explicitly told about the false memory effect.

Results revealed that recognition for studied items was highest for the uninformed group, whereas the recognition of studied items was equivalent for the cautious and forewarned groups. Perhaps more importantly, the forewarned group falsely recognized fewer critical lures than both uninformed and cautious groups (Gallo et al., 1997), although they still falsely recognized over half of all critical lures. The results from McDermott and Roediger (1997) and Gallo et al. (1997) provide converging evidence that although the false memory effect is weakened when participants are warned about the false memory effect, participants still reliably report recognizing critical lures that were not presented on the word lists.

Fuzzy trace theory would predict this result if one assumes that when individuals are warned of the effect, they automatically adopt a verbatim strategy over a gist-based strategy for recall/recognition. The source monitoring theory would also predict this result if one assumes that when participants are warned, they adopt a higher threshold on which they base their memory judgments. Similarly, the criterion shift theory would predict this result if one assumes that participants adopt a higher criterion at recall/recognition. The spreading activation theory, on the other hand, could not predict this result.

The explanation of the false memory phenomenon was also used to directly test if the criterion shift account could be used to explain false memories. Gallo, Roediger, & McDermott (2001) wanted to determine if the warning effects could be due to encoding processes or decision processes at test (criterion shift) or a combination of both. Participants were place in one of four groups. The control group was instructed to be as accurate as possible (no warning condition). The second group was given the warning before the study, replicating the two previous studies mentioned (warning before study). The third group was given the warning after the words were presented but before the test (warning after study). These three groups were administered 12 lists from Roediger and

McDermott (1995) in which half the lists included the critical lure in the list in an attempt to manipulate strategies instead of calculating signal detection estimates of bias. One addition group was included in which the participants were given a warning after studying the 12 without any of the critical lures present (warning after study with all critical lures). After the presentation of the lists, all participants were administered the same 84-item recognition test and were instructed to make the Remember/Know distinction. All groups were given an example using the list 'sleep' which was used to explain the false memory phenomenon in the experimental groups.

Results revealed that the rate of critical lures reported as old was significantly higher when the corresponding list had been studied compared to when the list had not been studied for the control, warning before study and warning before test groups. There was no significant difference between the rate of critical lures identified as old in the warning before study group and warning before test group. The authors then employed several statistical corrections of the recognition data for both studied items and critical lures based on procedures by Snodgrass & Corwin (1988) with the purpose of providing an index of the influence of list presentation on recognition performance (see Gallo et al, 2001). This correction indicated no difference between the control, warning before study and warning before test groups in the recognition of studied items and critical items when the critical lure was presented. When the critical lure was not presented in the list, there was a significant difference between recognition of studied items and critical lures, such that studied items were reported as old at a higher rate. In addition, there was no difference between those three groups in the rate of studied items. However, the warning before study group significantly identified critical lures as old at a lower rate than the control group. The warning before test did not differ in the rate of false recognition of critical lures from the control group. The manipulation check, the warning before test group with all critical lures presented, indicated a response bias, such that they rated both critical lures and studied items as new more often than the warning before test group.

The results from this study indicate that presenting the critical lure in some lists but not in others prevents participants from adopting the criterion shift as a strategy, as the warning attempts to motivate the participants to use a stringent criterion. However, because the participants still showed a high rate of false recognition of critical lures, it is apparent they were unable to do so. Therefore, it is highly unlikely that the false memory phenomenon is simply due to a criterion shift in strategy. A theoretical account of the limitations of the criterion-shift account of false memory is provided mathematically in Wixted & Stretch (2000).

List Length

Deese (1959) proposed that false recall and false recognition would be related to the mean (average) associative strength of all the words on each list. The mean or average associative strength of a list is calculated by adding the strength of the each word’s association to the critical lure and dividing that sum by the number of words on the list. This hypothesis stands in contrast to the hypothesis of Robinson and Roediger (1997) in which they argue that false recall and false recognition is related to the total associative strength of all the words on the list. The total associative strength of a list is calculated by obtaining a sum of the strength of each word's association to the critical lure.

Using the Roediger and McDermott (1995) paradigm, Robinson and Roediger (1997) examined these two hypotheses by changing the total number of associated items in each list (list length). In their experiment, five different list lengths were presented with 3, 6, 9, or 15 words in each list. In each list, the most highly associated word to the critical lure was in the first position and the least associated word was placed in position 15. Thus, the different lengths included a different number of the most highly associated words to a critical lure in an attempt to manipulate the total associative strength of each list. Results revealed that participants recalled fewer list words but more critical lures as list length increased. Surprisingly, increasing the list length had no effect on correct recognition, but did increase the false recognition of critical lures. On the basis of this evidence, Robinson and Roediger (1997) concluded that the magnitude of the false memory effect increased with total associative strength.

In a second experiment, Robinson and Roediger (1997) sought to replicate this finding using the same word lists, but they added unrelated words so that each list contained 15 words. If the mean (average) associative strength of the list has an effect on false recall and recognition as Deese (1959) proposed, then the results from Experiment 2 should not replicate those obtained in Experiment 1. In other words adding non-associated words to equalize list length would diminish the mean associative strength of the lists in Experiment 2. However, if the total associative strength has an effect on false recall and recognition as Robinson and Roediger (1997) suggest, then the results from Experiment 2 should replicate those obtained in Experiment 1 because the total associative strength is equivalent between the lists in the two experiments. Results from Experiment 2 replicated those obtained in Experiment 1, confirming Robinson and

Roediger's (1997) hypothesis that it is the total associative strength of the words in the list that influences false recall and recognition.

The results obtained from Robinson and Roediger (1997) can be readily explained by the implicit and spreading activation theories. With respect to the implicit activation theory, as the number of associated words in the list increases, the number of times the critical lure is activated should also increase. The more the critical lure is activated by associated words, the more likely it is that a participant will implicitly activate the critical lure. This activation should not depend on how many words are on a list, but only on how many words on the list are associated with the critical lure. The same logic may be applied to show how the semantic activation theory could predict this result. The source monitoring theory fuzzy trace theory, and the criterion shift theory, as currently proposed, do not make clear predictions about the mean associative versus total associative strength issue.

Encoding and Retrieval

Attempts have been made to study the effect of manipulating the encoding and retrieval of the words on each list. In Experiment 2, Read (1996) placed participants in three conditions: a serial-learning group, in which participants were told to learn the words in the order in which they are presented; an elaborative-rehearsal group, in which participants were instructed to think about and rehearse the words in order to answer questions about word meanings later; and a maintenance-rehearsal group, in which participants were told only to remember the last word of each list. At recall, participants were told either to write down as many words as they could in free-recall format or to recall the words in the order in which they appeared. Results showed that the elaborative-rehearsal group recalled more list words than the maintenance-rehearsal group who, in turn, recalled more list words than serial-learning group. With regard to critical lures, the serial-learning group recalled significantly fewer critical lures than the elaborative- and maintenance-rehearsal groups, who did not differ from each other with respect to the number of critical lures recalled.

The results obtained by Read (1996) are difficult to explain by any of the proposed theories of false-memory construction. Perhaps the best explanation can be offered by the fuzzy trace theory. The serial learning group presumably processed the words in the lists most deeply (Craik & Lockhart, 1972), and relied more on verbatim than gist traces which would explain why they recalled fewer critical lures than the elaborative- and maintenance-rehearsal groups. Because the latter two groups were not required to remember the words in order, they could rely on both verbatim and gist traces which would increase the probability that they would recall more critical lures. In contrast, the spreading activation theory, the implicit activation theory, the source monitoring theory and the criterion shift theory do not easily explain the results obtained by Read (1996).

Social Factors

Recently, there has been some work examining the effects of social factors on false memories. Roediger, Meade and Bergman (2001) combined two famous paradigms in studying memory: Asch's (1952) paradigm to study social conformity and Loftus' (1993) paradigm studying eyewitness testimony. Six slides consisting of common household items were developed for this experiment with each slide containing items that participants in a pilot study indicated they would expect to see. Items were classified as either high- or low-expectancy items based upon the number of participants indicated they expected to see that item. Four items were excluded from each slide, 2 low- and 2 high-expectancy items, to be used later in the experiment. Participants were accompanied by a confederate. The participant and the confederate were shown the 6 slides for either 15 seconds or 60 seconds. After all 6 slides had been presented, multiplication problems were administered for 4 minutes. The participant and the confederate were then asked to work together on a free recall test in which they took turns saying one item they remembered for a total of 12 items for each of the 6 slides. The confederate was instructed on 3 of the slides to produce one high- and one low-expectancy (contagion) item that had been omitted in the slides. The participant was then taken into a separate room and completed a free recall test for each scene with additional instructions to indicate if they remembered the item or if they knew the item had been seen.

The results indicated that the participants falsely recalled the contagion items suggested by the confederate at a higher rate than the contagion items not suggested by the confederate. Participants were also more likely to recall high-expectancy items than low-expectancy items. In addition, participants who were only allowed to view the scenes for 15 sec were more likely to include the suggested items than participants who were allowed to view the scenes for 60 sec. The remember/know responses indicated that the contagion items which were falsely recalled were more likely to be judged as known than remember.

The source monitoring theory provides the best account for the results. According to the source monitoring theory, the collaborative recall test is a source of retroactive interference whereby the participants later retrieve the erroneous information presented during the collaborative recall, but attribute it to the presented scene rather than the confederate. Therefore, the participant would falsely recall items recalled by the confederate. In addition, the know judgments indicate that although the item is attributed to the scene, the exact source is unavailable to the participant.

The spreading activation theory and the fuzzy trace theory would also be able to explain the results, although on a much more basic level. The spreading activation theory would propose that the items the confederate suggested during recall had already been activated by the presentation of the scene, and were pushed beyond threshold by the confederate's mention of the items. The participant would recall the items as if they had actually been presented during the slides, but the actual memory would be unavailable and those items would be judged as known. The explanation through the fuzzy-trace theory is similar, only the participants do not rely on activation of the nodes, they rely on fuzzy-traces, some of which would lead to those items mentioned by the confederate.

Development

There have been several studies that have examined false memories in children and older adults. In Underwood's (1975) terms, the examination of age-related differences may serve as a 'crucible' for evaluating more general theories of cognitive function.

A multitude of studies have examined adult age differences in younger and older adult's susceptibility to false memories (e.g., Norman & Schacter, 1997; Tun, Wingfield, Rosen, 1995; Tun, Wingfield, Blanchard, Rosen, 1996). In one experiment, Norman and Schacter (1997) tested both young and older adults using the Roediger and McDermott

(1995) paradigm. In addition, they modified the recognition test such that on items that were recognized as old, participants were not only asked to make remember and know judgments, but they were also asked to explain why they made the judgments that they did (e.g. I reported that “needle” was old because it was in the same list as thread).

Older adults recalled fewer studied items, more critical lures, and more intrusions (novel items) than younger adults. On the recognition test, older adults recognized fewer studied items, more critical lures, and falsely identified more new items as having been on the list as compared to young adults.   Although more ‘know’ responses were generated for critical lures when both young and old participants were required to explain their responses, participants still falsely recognized critical lures. Together, these findings indicate that there are developmental differences in false-memory construction favoring young adults, and that even careful consideration of critical lures did not diminish the false memory phenomenon.

Analysis of the reasons why participants reported some words as old revealed that the information about the semantic relationships between words on the list was important for both studied items and critical lures, regardless of the age of the participant. Young and old participants also reported that the reaction(s) that they had to a word presented on the list helped them to identify words on the recognition test as old (e.g. I remember the word arachnid because of the fear I have of spiders). This indicates that both younger and older adults’ explanations were primarily associative in nature.

The results obtained by Norman and Schacter (1997) may best be accounted for by the fuzzy trace theory of memory. Because older adults have poorer memory abilities, they are likely to rely more on the fuzzy trace information rather than the verbatim information, making them more susceptible to false-memory construction than young adults, who have more verbatim information available to them. However, the criterion shift and spreading activation theories could also explain the difference between younger and older adults if one assumes that older adults adopt a more lenient criterion (or lower threshold) when making judgments on items because of reductions in their memory ability.

Recall Confidence

Researchers in false-memory construction have been interested in the confidence with which participants recall information from words on lists that were presented to them. It may be that participants recall or recognize critical lures in an experimental setting, but are not very confident that they heard or saw those words on the studied lists. To investigate this possibility, Read (1996, Experiment 1) conducted an experiment using the Roediger and McDermott (1995) paradigm in which participants were required to give a confidence rating for each word that they reported during a recall test. Participants were to assign a confidence rating from 1 (no confidence) to 5 (extremely confident) as a testament to how confident they were that word was presented. Results revealed that when participants reported the critical lure early in recall, it received a similar confidence rating as those given for studied items. However, when participants reported the critical lure late in recall, it received a lower confidence rating as those given for studied items. This finding suggested that participants' subjective confidence in information recalled from memory is not necessarily correlated with the accuracy of that information (Read, 1996).

A study by Brédart (1999) employed the confidence intervals used by Read for the recall phase, and additionally examined why some participants did not recall the critical lure. The author proposed two hypotheses: the critical lure was not thought of at all, or else it was thought of, but the participants determined that the word had not been presented, and therefore did not write it down. Eight lists of ten items were constructed based on two pilot studies that determined degree of association for those lists and their ability to generate a mental image of the person/character associated with the lists. Each list was read out to the participants, and then followed by recall for 90 sec (Phase I). In Phase II, participants were asked to rate on a scale from 1 to 5 how confident they were, that each word they recalled had been read aloud by the examiner. In addition, in Phase III participants were presented with the words they recalled for each list separately, and were asked if there were any words they had thought of for that list either during the presentation or recall, but had not written down because they thought it had not been read aloud. Participants were then asked to rate their confidence on a scale from 1 to 5 how sure they were, that the word had not been presented to them.

The results indicated that there was no difference in the amount of critical lures recalled in Phase I and Phase III. However, in Phase III, a comparison between participants who recalled critical lures in Phase III and participants who did not recall the critical lure during Phase III were compared for each list. The results from this analysis revealed that for each list, more participants recalled the critical lure at a significantly higher rate than not. The author concluded that this was evidence that participants did not include the critical lure during Phase I recall because, although they thought of the critical lure, they were able to determine that it had not been presented.

This is an example of successful source monitoring. The critical lure was activated at or above threshold, according to the spreading activation theory and the implicit activation theory, because the critical lure was recalled during Phase III, but neither the spreading activation theory nor the implicit activation theory would be able to explain why the critical lure was not recalled during Phase I. On the other hand, the fuzzy-trace theory would be able to explain the results, as the participants were basing their judgments more on verbatim traces than fuzzy traces in Phase I, but were asked to utilize fuzzy traces in Phase III, thereby producing the critical lure.

Metamemory

In contrast to research that has examined participant's confidence in their recall of words presented on lists, it is possible that one's overall confidence in one's memory, or metamemory, might affect false-memory construction. Several studies have shown that memory confidence is positively related to recall and recognition performance (e.g., Dixon & Hultsch, 1983b, Williams, Denney & Schadler, 1983). In other words, participants who believe that they have good memory abilities also tend to show higher levels of performance on recall and recognition tests than participants who believe that they have poor memory abilities. If memory confidence affects recall and recognition performance, then it stands to reason that memory confidence might also affect false-memory construction, although the direction of this relationship is unclear.

Memory Confidence and Eyewitness Testimony

One of the factors in eyewitness testimony that jurors rely heavily upon is the confidence of the eyewitness in their ability. Common sense would identify an eyewitness's confidence in their ability to identify the criminal would help indicate how accurate an eyewitness is. In fact, the court case Neil vs. Biggers (1972) identified confidence as an index of reliability and accuracy of the eyewitness. Wells, Lindsay and Ferguson (1979) were able to establish that not only judges, but also juries, place importance on the confidence of the eyewitness. Migueles and Garcia-Bajos (1999) investigated confidence in free recall and recognition tests in eyewitness testimony under certain conditions. Participants viewed three video segments: the first segment contained news about drugs, the second was an emotional scene of a kidnapping attempt, and the third contained commercials. One group of participants (incidental group) was told they would be asked to judge the length of each of the segments while the other group (intentional group) was told they were told to pay close attention as they would be asked to make evaluations about the segments later. Participants were given five minutes after viewing the segments to recall what took place in the kidnapping scene, and were then given a recognition test on which they evaluated sentences by determining if each sentence was true of false, and their confidence of their decisions. The results concerning their confidence ratings revealed that confidence ratings were higher when participants correctly identified a statement was true than when they incorrectly identified the statement was true, which in turn was higher then the ratings on correct rejections and incorrect rejections. Participants also had higher confidence scores with questions concerning central information than peripheral information, and higher confidence scores on action questions than details. In conclusion, confidence scores depended on the type and content of the information as well as the accuracy.

Recently, Schneider, Jenkins, and Ciampi (unpublished data examined the extent to which memory confidence influenced recall, recognition, and false memory performance using the Roediger' and McDermott (1995) paradigm by manipulating participant's confidence in their memory abilities experimentally. Specifically, participants were given false feedback about their performance on a memory game (concentration). Some participants were told they had performed poorly on the memory game (low-confidence group), others were told that they had performed extremely well on the memory game (high-confidence group), and the remaining participants were given no feedback about their performance (neutral group). After the instructional manipulation was administered, the Roediger and McDermott (1995) false memory paradigm was administered to all participants.

Results failed to reveal a significant condition effect recall or recognition of critical lures. However, there was a trend in the recall data for studied words, such that the high-confidence group recalled studied words at a rate of .57, the control group recalled studied words at a rate of .55, and the low-confidence group recalled words at a rate of .52. These findings suggest the possibility that the confidence manipulation was not strong enough to influence recall or recognition performance.

Present Study

Although no effect of memory confidence was obtained in the Schneider et al. study, two mitigating factors need to be considered before one can conclude that metamemory does not affect false-memory construction. First, Schneider et al. used a between subjects design to investigate their hypothesis; an experimental design that is considerably weaker than a within subjects design. Second, the lack of an effect of metamemory on the recall and recognition of studied items and critical lures may have resulted from an unsuccessful or weak confidence manipulation.

In light of these limitations, the goal of the proposed study was to further investigate the effects of metamemory on false-memory construction. This research improves on Schneider et al. by manipulating memory confidence within-subjects. Additionally, the present study used performance on three memory tests in an attempt to increase the strength of the metamemory manipulation. Initially, participants were randomly assigned to one of three groups: high confidence, low confidence, or neutral. Participants in the high-confidence group received positive, negative, and then positive feedback about their memory performance on three different memory tests. Participants in the low-confidence group received negative, positive, and then negative feedback about their memory performance on the same three memory tests. Finally, participants in the neutral group received neutral feedback about their memory performance on all three memory tests.

Using the Roediger and McDermott (1995) paradigm, participants were administered a total of 32 word lists taken from Stadler, Roediger, and McDermott (1999), and a recall test was administered after each list was presented. Word lists were divided into 4 separate blocks of 8 lists. After each block of lists, participants took one of the three memory tests. Participants in each group received the appropriate feedback about their performance in order to manipulate metamemory in the desired direction (positive, negative, or neutral). After all 32 lists were presented, participants completed a recognition test in which they were asked to identify the words presented on each list and to make remember, know (Tulving, 1985), and guess judgments (Gardiner & Conway, 1999)

Hypotheses

Hypothesis 1

As in previous research (e.g., Roediger & McDermott, 1995), participants will show higher rates of correct recall for studied items as compared to critical lures. In addition, participants will show higher rates of false recall of critical lures than non-studied intrusions.

Hypothesis 2

If memory confidence affects recall, then the recall rate of studied items will change after each manipulation for both the high and low confidence groups. That is, the high confidence group will recall more studied items from lists 9-16 and 25-32 than from the lists 1-8 and 17-24, and will recall fewer studied items from lists 17-24 than any other list. The low confidence group will recall more studied items on lists 17-24 and than any other list. The low confidence group will also recall fewer studied items on lists 9-16 and 25-32 than from lists 1-8 and 17-24. The neutral group will not change in the amount of studied items recalled on any of the lists. These predictions are displayed in the three panels of Figure 2.

If memory confidence affects recall, then it is anticipated that memory confidence will also have an effect on the recall of critical lures. However, because of the exploratory nature of the proposed study and because no one theory of false memory construction can account for all the results obtained in previous research, no formal predictions about the nature of this relationship are offered.

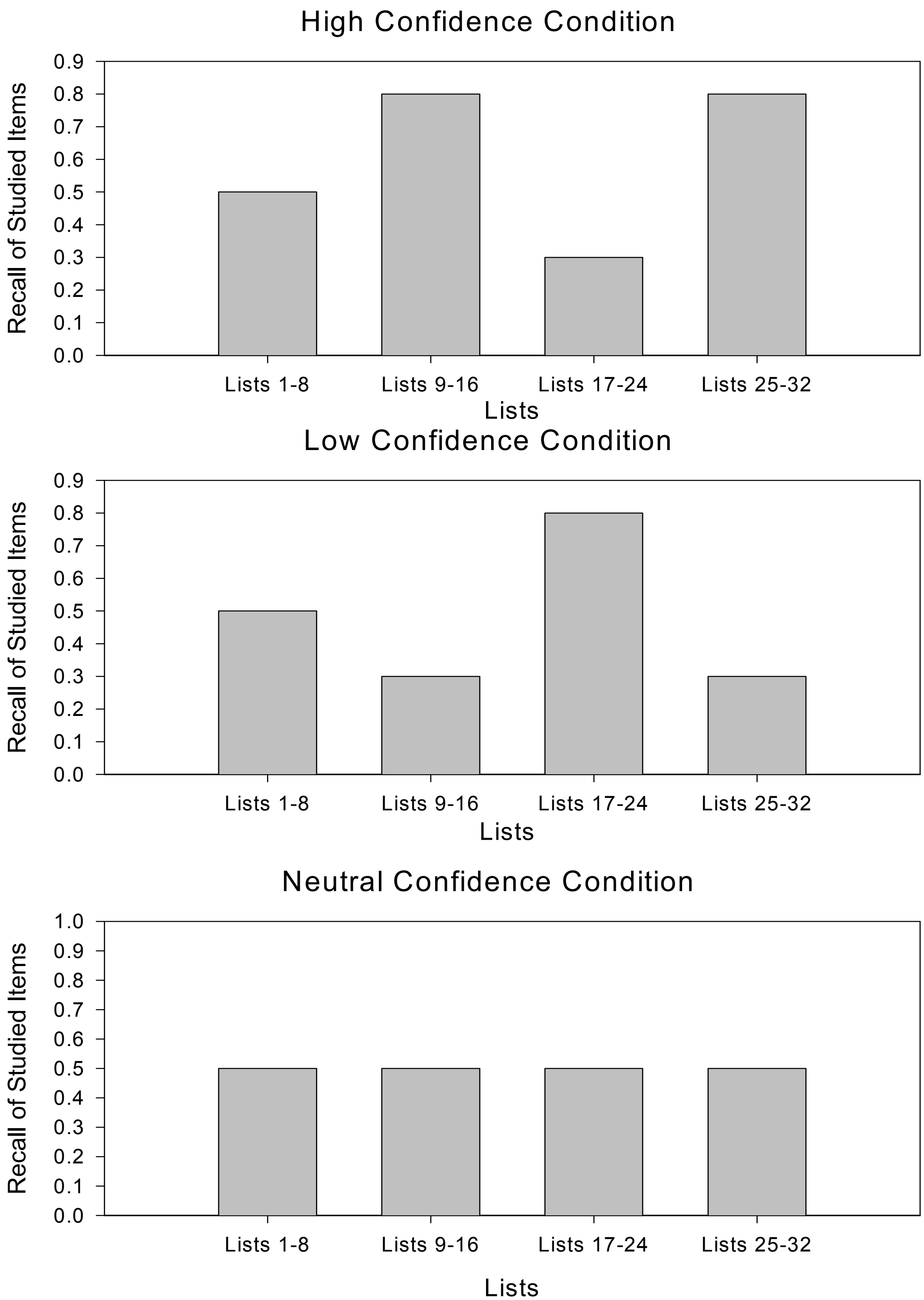


Figure 2. Hypothesized results for the recall of studied items for the high-confidence group (top), the low-confidence group (middle), and the neutral group (bottom).

Hypothesis 3

As in previous research (e.g., Roediger & McDermott, 1995), participants will show higher rates of correct recognition for studied items as compared to critical lures. In addition, participants will show higher rates of false recognition of critical lures than non-studied items.

Hypothesis 4

If memory confidence affects recognition, then the recognition of studied items and critical lures will change after each manipulation for both the high and low confidence groups. The high confidence group will recognize more studied items from lists 9-16 and 25-32 than from the lists 1-8 and 17-24, and will recognize fewer studied items from lists 17-24 than any other list. The low confidence will recognize more studied items from lists 17-24 and than any other list. The low confidence group will also recognize fewer studied items from lists 9-16 and 25-32 than from lists 1-8 and 17-24. The neutral group will not change in the amount of studied items recognized on any of the lists. These predictions are displayed in the three panels of Figure 3.

If memory confidence affects recognition, then it is anticipated that memory confidence will also have an effect on the recognition of critical lures. However, because of the exploratory nature of the proposed study and because no one theory of false memory construction can account for all the results obtained in previous research, no formal predictions about the nature of this relationship are offered.

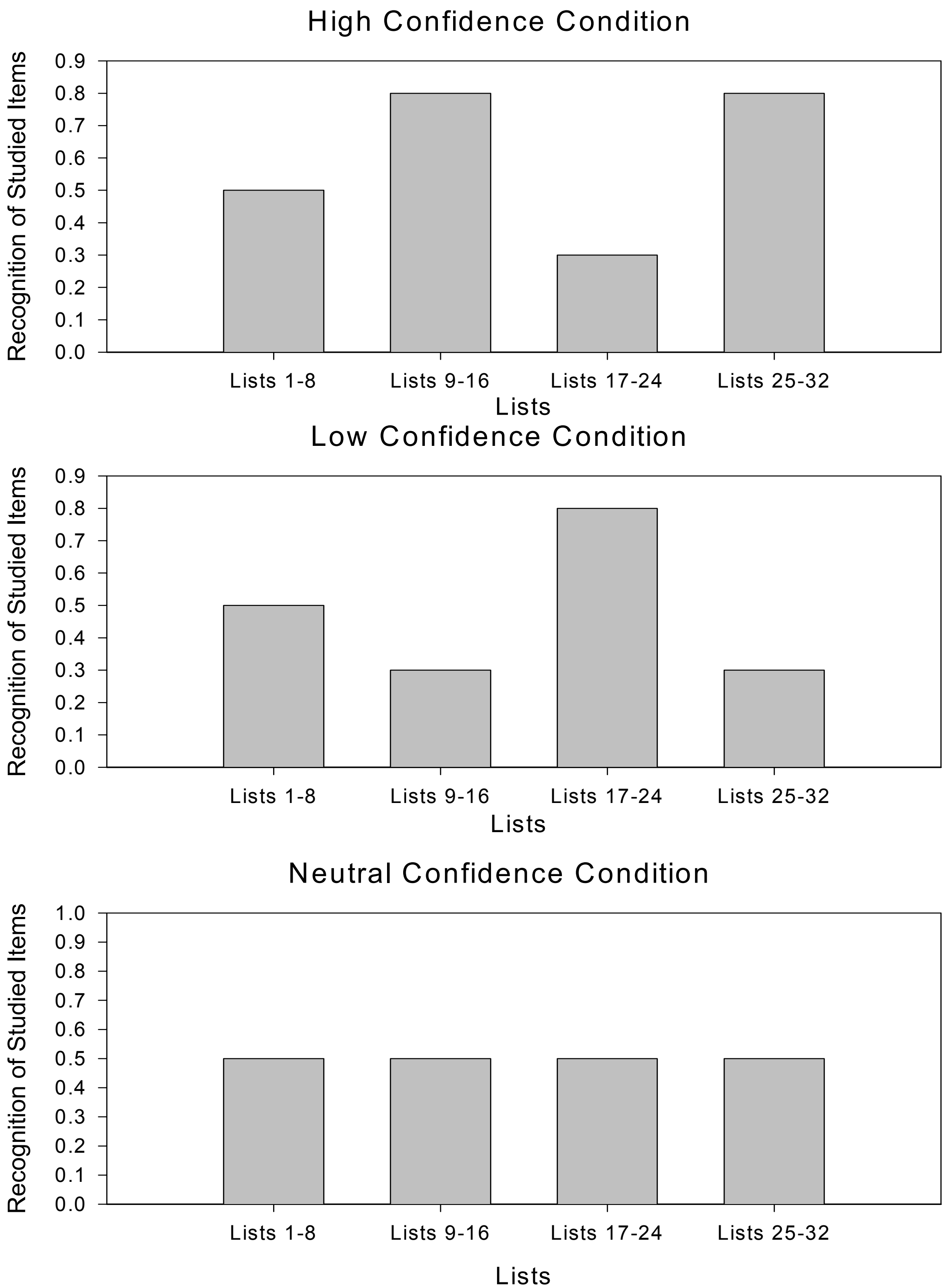


Figure 3. Hypothesized results for the recognition of studied items for the high-confidence group (top panel), the low-confidence group (middle panel) and neutral group (bottom panel).

Hypothesis 5

Consistent with results obtained by Roediger and McDermott (1995), participants will make more remember judgment on studied items recognized as old, than critical lures recognized as old. In addition, more know responses will be made on critical lures identified as old than studied items recognized as old. It is possible that the generation of high memory confidence may result in participants making more remember than know responses whereas the generation of low confidence may result in participants making more know than remember responses, although these predictions are purely speculative in nature.

METHODS

Participants

Forty-three females and 17 males aged 18-25 (M = 19.08, SD = 2.06) participated, yielding an overall sample of N=60 participants. Participants were recruited from the University of North Carolina Department of Psychology undergraduate General Psychology volunteer subject pool. Each participant was screened using a detailed self-report health questionnaire (see Appendix A) and participants were excluded if they reported a history of alcohol or drug abuse, drug/alcohol addiction, hospitalization within the last year for a serious medical illness, specific visual problems (e.g., color blindness), medical problems (e.g., diabetes), neurological problems (e.g., traumatic head injury) or psychiatric illnesses (e.g., current or chronic depression). Information on the use of over-the-counter and prescription medications was obtained and participants included in the study were not taking any medications known to affect general arousal, motivation, memory, or learning.

Analyses were performed on the age, sex and education of the participants to determine whether there were significant differences between the confidence groups. There were significantly more females than males in the neutral confidence condition, $\chi^2(1, N = 20) = 7.2, p < .05$ than in the high and low confidence conditions, but no gender imbalances were observed between the high and low confidence conditions. As there is no literature to suggest gender differences in recall or recognition performance on studied items or critical lures, the finding of a gender difference between the groups is not expected to systematically alter the results of this study.

One-way ANOVAs were performed on the age and education of the participants. Results revealed there no differences in the ages (F = .23, MSE = 4.36) or years of education (F = .30, MSE = 1.5) of the participants in the different conditions (ps > .05).

A 8 (Lists: 1-8) x 3 (Condition: High, Low, Neutral) ANOVA was performed on the recall of the first 8 lists was performed to examine whether there were any overall differences memory performance on recall between the conditions. The ANOVA revealed no significant main effects ($p < .001$), or interactions ($p > .05$).

	Groups			
	Overall	High	Low	Neutral
Age	19.08 (2.06)	18.85 (1.46)	19.1 (1.74)	19.3 (2.81)
Education	12.63 (1.21)	12.78 (1.30)	12.48 (.99)	12.65 (1.35)

Table 1. Means (Standard Errors) of Age and Education by Group

Materials

Health Questionnaire

The self-report health questionnaire (Appendix A) was designed to assess current sensory motor functioning, prior history of medical illness or disease, and prior history of possible psychiatric illness (Jenkins, Myerson, Joerding, & Hale, 2000). The health questionnaire was administered for two purposes. First, responses to the health questionnaire were used for screening purposes to ensure that a healthy, homogenous sample was obtained. Second, responses on the health questionnaire were used to assess the existence of baseline differences in motor, visual, or psychological function between individuals assigned to the high, low, and neutral-confidence groups.

Pre-Test Memory Questionnaire

This questionnaire (Appendix B) consisted of 3 questions designed to assess participant's perceptions of their own memory abilities. The primary purpose of this questionnaire was to evaluate the efficacy of the confidence manipulation at the end of the study.

Word Lists

A total of 48 words lists (Appendix C) were used in the present study: 36 word lists taken from Stadler et al. (1999) and 16 lists were constructed using the same norms Stadler et al (1999) used from Russell and Jenkins (1954). Each list consisted of 15 words highly related to one non-presented word (critical lure). Only 32 lists were presented to the participants, and the remainder was used for the recognition test. The presentation of the words in each list was randomized, but all participants received the same random ordering of words in each list. The number of words correctly recalled, the number of critical lures recalled, and the number of intrusions (i.e., words recalled that are not on the word lists or are critical lures) were measured. The mean associate strength of each list based upon Roediger, Watson, McDermott, and Gallo (2001) was included in Appendix C if the information was available for that list.

Rey-Osterrieth Complex Figure Test (Rey, 1944)

This test (Appendix D) was used to assess the visual-memory ability of the participants as well as provide the basis of the first memory manipulation. The Rey-Osterrieth Complex Figure test was comprised of 18 interconnected elements (lines, squares, and circles). The figure was given to the participants to study for a total of 2 minutes and then taken away. The participant was given a blank piece of paper and told to reconstruct the figure as they remember it. Participants were given 5 minutes to complete their drawing. This test was scored using Loring, Martin, Meador, and Lee's (1990) scoring system. They developed a scoring system that placed more emphasis on memory rather than constructional elements. Each of the 18 elements was scored between .5-2.0 depending accuracy, distortion, and location. Two points were awarded if the participant correctly replicated the element and placed in the proper position. One point was awarded if the element was correct, but not placed properly. If the element drawn was distorted and placed poorly, one-half point was awarded. The highest score possible was 26.

Intense Concentration

This test was a version of the childhood memory game, "Concentration" that was computerized by Michael Hawden for use in the Schneider, Jenkins and Ciampi (unpublished data) study and the present experiment. In this test, the computer presented a display of sixteen blank cards (1.5 in x 1.5 in) that were arranged in a square with 4 rows and 4 columns. Participants were instructed to use a mouse to click on any card in the matrix to reveal a picture that appears on the other side. Participants were told that there are eight unique pictures and that their job was to find the match for the picture on each card as quickly as possible. Participants were also told that they may only turn over two cards at a time. In a given turn, if the two pictures that they turn over were identical, then both pictures will remain revealed. However, if the two pictures did not match, then the pictures disappeared and both cards appeared blank again. The game was completed when all squares were revealed. There was no time limit to complete this task, although scores were computed by the game for both the time it took the participant to win the game, and the number of times the participant had to click on the squares before completing the game. The pictures used in this game consist of light purple donut- and bubble-like shapes that were presented on a dark purple background. The pictures were randomly chosen from the set of BubbleMania Soft 2C fills in Corel Draw 9.

WAIS-R Digit Span (Wechsler, 1981)

In this test (Appendix E), participants were read a series of numbers one at a time with a one second break between each presentation. In the forward digit span task, participants were to report the numbers in each series in the order that they were presented. For example, if the experimenter said "6, 8, 1", the participant was to say "6, 8, 1". In the backward digit span task, participants were to report the numbers in each series backward. For example, if the experimenter said "6, 8, 1", the participant is to say "1, 8, 6". In both tasks, participants were exposed to two sets of series at an initial length of 2 items. If at least one of the series was recalled correctly, the length of the series was increased by one. The experimenter discontinued testing when the participant failed to correctly recall two series of the same length. Forward and backward memory spans were calculated separately. One point was given for every correctly recalled series length and an individual's memory span in each task was the sum total of all series correctly recalled. No penalty was given for errors and no time limit was placed on this task.

Digit Symbol Substitution Test (Wechsler, 1981)

This test (Appendix F) was designed to examine processing speed and working memory ability. At the top of the test, a key was presented in which the numbers 1- 9 are each associated with one symbol that appears in a box below each number. At the bottom of the page, seven rows of 20 numbers (ranging from 1-9) were presented with a blank box beneath each number. Participants were told to place the corresponding symbol beneath each number in the box provided, starting at the top left hand corner and proceeding across to right. Seven sample items were given, after which the participant had 60 seconds to complete as many of the boxes that they could. Participants were awarded one point for every correct symbol, and no penalties were given for incorrect or absent symbols. A total of 133 points could be obtained on this test.

Recognition Test

In this test (Appendix G), words that were presented in positions 1, 8 and 10 and all 32 critical lures from the word lists were presented. In addition, words in position 1, 8, and 10 and the critical lure from the 16 word lists not presented were also included on the recognition test as novel words. In sum, there were 192 words on the recognition test (96 studied list items, 32 critical lures, and 64 non-studied or novel words). The presentation of words on the recognition test was randomized with the constraint that no words or critical lures from the same list were presented consecutively. All participants were administered the same randomized order of words on the recognition test.

Participants were to indicate whether they recognize the word as one that they studied by circling “old” or as one that is novel by circling “new.” When a word on the recognition test was identified a studied item, participants were also to indicate whether they specifically remembered the word being presented (remember) by writing ‘R’ on the line provided, simply know that the word was presented (know) by writing ‘K’ on the line provided, or guessed that the word was on a list (guess) by writing ‘G’ on the line provided. Remember and know judgments were included in order to determine if participants who recognized critical lures as being presented actually report remembering the item or not. Guess judgments were also included for those instances when the participant could not remember the item as having been presented and has no feelings of knowing that the item has been presented, thereby increasing construct validity (Gallo, McDermott et al, 2001)

The number of studied words correctly recognized as old (i.e., hits) and the number of novel words correctly identified as new (i.e., correct rejections) were measured. In addition, the number of critical lures identified as old and the number of novel words incorrectly identified as old (i.e., false alarms) were measured. Finally, the number of studied words identified as new (i.e., misses) was measured.

Post-Test Memory Questionnaire

This questionnaire (Appendix H) consisted of 4 questions designed to assess participant's perceptions of their own memory abilities. The data obtained from this questionnaire was compared to the data collected from the pre-test memory questionnaire in order to assess the global effects of the memory confidence manipulation on metamemory.

Procedure

Participants were tested individually in a quiet laboratory room and each testing session lasted approximately 1.5 - 2 hours. Participants were exposed to 11 tasks presented in the following order: pre-testing memory questionnaire, false memory lists 1-8 (including recall tests), Rey-Osterrieth Complex Figure test, false memory lists 9-16 (including recall tests), Intense Concentration, false memory lists 17-24 (including recall tests), WAIS-R Digit Span Subtest, false memory lists 25-32 (including recall tests), WAIS-R Digit Symbol Substitution Test, the false memory recognition test, and the post-testing memory questionnaire.

Initially, participants were administered the health questionnaire and the pre-testing memory questionnaire. Both questionnaires were self-paced. After completing the pre-testing memory questionnaire, participants were administered the first false memory task (lists 1-8). Participants were told that they would see a series of words on the computer screen presented one at a time. Each word was presented on the screen for 1 s with a blank screen appearing for 500 ms before the appearance of the next word. After each word list was presented, a recall test was administered. Specifically, participants were given a blank piece of paper and told they had 2 min during which time they were to write down as many words from the list they just saw that they could remember. After the recall test for the eighth word list was completed, participants were given the Rey-Osterrieth Complex Figure test.

Upon completion of the Rey-Osterrieth Complex Figure test, all participants were told that this task taps memory abilities similar to those that they use when trying to find their car in a crowded parking lot (e.g., the mall). Participants were then quasi-randomly assigned to one of three groups (i.e., high confidence, low confidence, or neutral) with the constraint that there were equal numbers of participants assigned to each group. Participants in the low-confidence group were told that their memory performance on the Rey-Osterrieth Complex Figure test was below average. Further, they were asked if they have trouble finding their car in a crowded parking lot, and more generally, if they thought that their overall memory abilities were poor. Participants in the high-confidence group were told that their memory performance on the Rey-Osterrieth Complex Figure test was above average, suggesting that their overall memory abilities were exceptional. Finally, participants in the neutral group were told that their memory performance on the Rey-Osterrieth Complex Figure test was average, suggesting that their overall memory abilities was average.

Participants were then administered false memory lists 9-16. After completing the recall test from the sixteenth list, participants were administered the Intense Concentration Memory test. Participants were told that this task taps memory abilities similar to those that they use when trying to remember the name of an acquaintance or someone they just met.

Upon completion of the Intense Concentration Memory test, participants were given feedback about their performance on this test. Specifically, participants who were assigned to the low-confidence group were told that their memory performance on the Intense Concentration Memory test was above average, suggesting that, in contrast with the previous memory test, their overall memory abilities were excellent. Participants in the high-confidence group were told that their memory performance on this test was below average, suggesting that, in contrast with the previous memory test, their overall memory abilities were poor. Finally, participants in the neutral group were told that their memory performance on the test was average, suggesting that, consistent with the previous memory test, their overall memory abilities were average.

Participants were then administered false memory lists 17-24. After completing the recall test from the twenty-fourth list, participants were administered the WAIS-R Digit Span Subtest (Appendix E). Participants were told that this task taps memory abilities similar to those that they use when trying to hold a phone number in memory until they reach the phone.

Upon completion of the WAIS-R Digit Span Subtest, participants were given feedback about their performance on this test. Specifically, participants who were assigned to the low-confidence group were told that their memory performance on the WAIS-R Digit Span Subtest was below average, suggesting that, in contrast with the previous memory test, their overall memory abilities were poor. Participants in the high-confidence group were told that their memory performance on this test was above average, suggesting that, in contrast with the previous memory test, their overall memory abilities were excellent. Finally, participants in the neutral group were be told that their memory performance on the test was average, suggesting that, consistent with the previous memory test, their overall memory abilities were average.

Participants were then administered false memory lists 25-32. After the recall test from the thirty-second list was completed, participants were administered the WAIS-R

Digit-Symbol Substitution test (Appendix F). Immediately following the completion of the Digit-Symbol test, the recognition test was administered. Participants were asked to distinguish whether each word on the recognition test (Appendix G) was presented in any of the 32 lists (old) or if the word was novel (new). For each item judged to be old, participants were also be asked to report whether they remember the word (R), simply know that the word was presented in one of the lists (K), or if they guessed that the word was old (G). They were told to make a distinction for each item they recognize as old if they "remember its occurrence in the lists or whether they simply knew on some other basis that the item was a member of the study lists" (Tulving, 1985, p. 8) or if they just 'guess' the item had been presented to them. The participant was allowed to complete the recognition test at their own pace.

After participants completed the recognition test, they were administered the post-test memory questionnaire (Appendix H). This questionnaire was self-paced. Once the questionnaire was completed, the experimenter partially debriefed the participant as to the aims of this study and any questions were answered. Participants were only told about the false memory effect in order to minimize the possibility of the study being compromised. At the completion of the study, a letter was sent home to all participants explaining the true nature of the experiment.

RESULTS

All statistical tests were performed on PC SAS version 8.0. All graphical representations of data were generated using Sigma Plot version 8.0. For all statistical comparisons, the type III sums of squares were generated and the Bonferroni correction was applied in a groupwise manner to control the rate of false positive conclusions (type I error).

Recall

Words recalled by participants were categorized as studied items, critical lures, or non-critical intrusions. The number of studied items recalled in list 1 through list 8 were calculated and then analyzed in a 2 (List: 1-8) x 3 (Condition: Low, High, Neutral) repeated measures ANOVA to determine if there were any recall differences between the groups. No main effect of group was observed ($p > 0.05$), suggesting that the recall performance of the 3 groups was similar on the first 8 lists.

Probabilities of recall of the total number of studied items and non-critical intrusions were calculated across each list within a set, and then divided by 120 (the sum of words presented in each set). Critical lures were calculated in the same manner, except the sum was divided by 18. A 3 (Item: Studied, Critical Lure, Non-Critical Intrusion) x 4 (List: 1-8, 9-16, 17-024, 24-32) x 3 (Condition: High, Low, Neutral) repeated measures ANOVA was performed. As predicted, a main effect of item, $\underline{F}$ (2, 114) = 415.46, $\underline{p} < .0001$, $\underline{MSE}$ = 0.045, was observed. Planned comparisons revealed a higher probability of correct recall for studied items (M = .59, SD = .08) as compared to critical lures, (M = .37, SD = .16), $\underline{F}$ (1, 59) = 82.11, $\underline{p} < .0001$, $\underline{MSE}$ = 0.04. In addition, the probability of false recall of critical lures was higher than the probability of non-critical intrusions (M = .03, SD = .02), $\underline{F}$ (1, 59) = 297.13, $\underline{p}$ < .0001, $\underline{MSE}$ = 0.02. The recall probabilities are displayed in Figure 4 by item and condition.

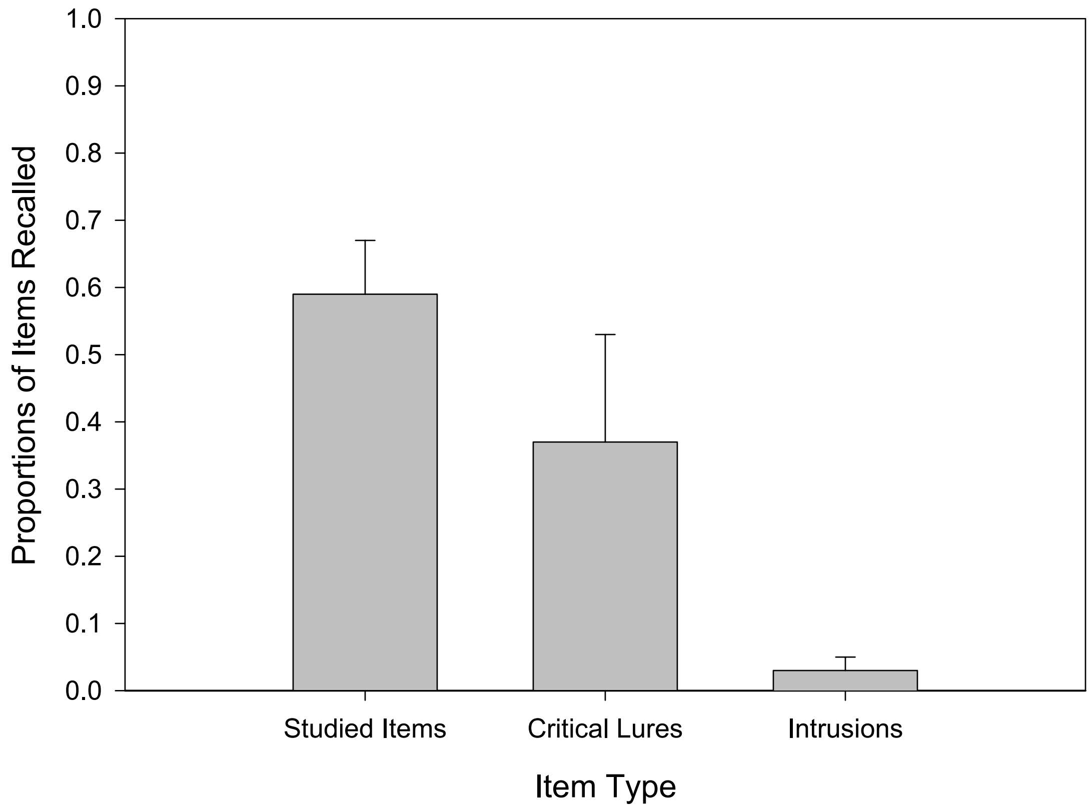


Figure 4. Probabilities of Studied Items, Critical Lures, and Non-Critical Items.

Recognition

Words recognized by participants were categorized as either studied items, critical lures, or non-studied items. The probability of recognition was calculated for all categories for all participants. Studied items were calculated first for each separate set of lists, and divided by 24 (the sum of studied items presented in each set). Critical lures were calculated in the same manner, only divided by 8. Non-presented items were calculated for the whole recognition test and divided by 64. A 2 (Item[1]: studied, critical lure) x 4 (List: 1-8, 9-16, 17-24, 24-32) x 3 (Condition: High, Low, Neutral) repeated measures ANOVA was performed and the proportions are presented in Table 2. The proportion of new items was excluded from analyses that included list as a factor because the new items could not be divided according to list presentation. As predicted, a main effect of item, $F(1, 57) = 19.91$, $p < .0001$, $MSE = .08$, as well as a main effect of list, $\underline{F}(3, 171) = 9.39$, $p < .001$, $\underline{MSE} = .02$, were found. The item by list interaction was also significant, $F(3,171) = 9.35$, $p < .0001$, $MSE = .01$.

Participants showed higher rates of correct recognition for studied items (M = .75, SD = .10) than rates of false recognition to critical lures (M = .37, SD = .19), $\underline{F}(1, 59) = 19.72$, $\underline{p} < .0001$, $\underline{MSE} = 0.042$. In comparison, participants showed higher rates of false recognition of critical lures (M = .62, SD = .19) than non-studied items (M = .12, SD = .06), $\underline{F}(1, 59) = 475.14$, $\underline{p} < .0001$, $\underline{MSE} = 0.029$.

Participants also showed higher rates of recognition of lists 1-8 (M = .70, SD = .16) than lists 9-16 (M = .63, SD = .13), $\underline{F}(1, 59) = 1638$, $\underline{p} < .0001$, $\underline{MSE} = 0.064$, lower rates of recognition on lists 9-16 than lists 17-24 (M = .72, SD = .14), $\underline{F}(1, 59) = 94.65$, $\underline{p}$

[1] New items were excluded from this analysis because they could not be divided into to list presentation groups.

< .0001, MSE = 0.057, and higher rates of recognition on lists 17-24 than on lists 25-32 (M = .68, SD = .15), F (1, 59) = 1805.51, p < .0001, MSE = 0.065. Further analyses indicated that lists 1-8 showed lower recognition than lists 17-24, F (1, 59) = 1740.81, p < .0001, MSE = 0.069. Participants also showed lower recognition on lists 9-16 than lists 25-32, F (1, 59) = 1698.44, p < .0001, MSE = 0.06.

The item by list interaction indicated that recognition of studied items changed between the lists. Participants recognized studied items at a lower rate on lists 1-8 (M = .71, SD = .13) than lists 9-16 (M = .72, SD = .13), F (1, 58) = 2304.80, p < .0001, MSE = 0.05. Participants recognized studied items at a lower rate on lists 9-16 than lists 17-24 (M = .80, SD = .12), F (1, 58) = 1679.18, p < .0001, MSE = 0.05. Participants recognized studied items at a higher rate on lists 17-24 than lists 25-32 (M = .72, SD = .13), F (1, 58) = 2789.59, p < .0001, MSE = 0.05. Further, participant recognized studied items at a lower rate on lists 1-8 than lists 17-24, F (1, 58) = 2725.19, p < .0001, MSE = 0.05. Participants also recognized studied items at a lower rate on listss9-16 than lists 17-24, F (1, 58) = 2479.09, p < .0001, MSE = 0.05.

Similarly, participants recognized critical lures at different rates across the lists. Participants recognized critical lures at a higher rate on lists 1-8 (M = .67, SD = .23) than lists 9-16 (M = .55, SD = .21), F (1, 58) = 562.5, p < .0001, MSE = 0.16. Participants recognized critical lures at a lower rate on lists 9-16 than lists 17-24 (M = .64, SD = .24), F (1, 58) = 493.75, p < .0001, MSE = 0.17. Participants recognized critical lures at a higher rate on lists 17-24 than lists 25-32 (M = .63, SD = .26), F (1, 58) = 493.75, p < .0001, MSE = 0.16. Further, participants recognized critical lure at a higher rate on lists 1-8 than on lists 17-24, F (1, 58) = 621.42, p < .0001, MSE = 0.17. Finally, participants recognized critical lure at a lower rate on lists 9-16 than on lists 25-32, F (1, 58) = 430.97, $p < .0001$, $MSE = 0.19$.

Lists	Overall	High	Low	Neutral
Correct Recognition of Studied Items				
1-8	.71 (.13)	.72 (.12)	.71 (.14)	.71 (.11)
9-16	.72 (.13)	.72 (.12)	.71 (.15)	.74 (.12)
17-24	.8 (.12)	.81 (.13)	.79 (.13)	.79 (.11)
25-32	.76 (.14)	.80 (.13)	.76 (.14)	.72 (.12)
Total	.75 (.10)	.76 (.10)	.74 (.11)	.74 (.09)
False Recognition of Critical Lures				
1-8	.67 (.23)	.68 (.23)	.71 (.23)	.61 (.23)
9-16	.55 (.21)	.55 (.22)	.60 (.19)	.51 (.22)
17-24	.64 (.24)	.58 (.23)	.71 (.28)	.63 (.20)
25-32	.63 (.26)	.56 (.27)	.71 (.23)	.60 (.28)
Total	.62 (.19)	.60 (.20)	.67 (.20)	.59 (.18)

Table 2. Means (Standard Errors) of Probability of Recognition.

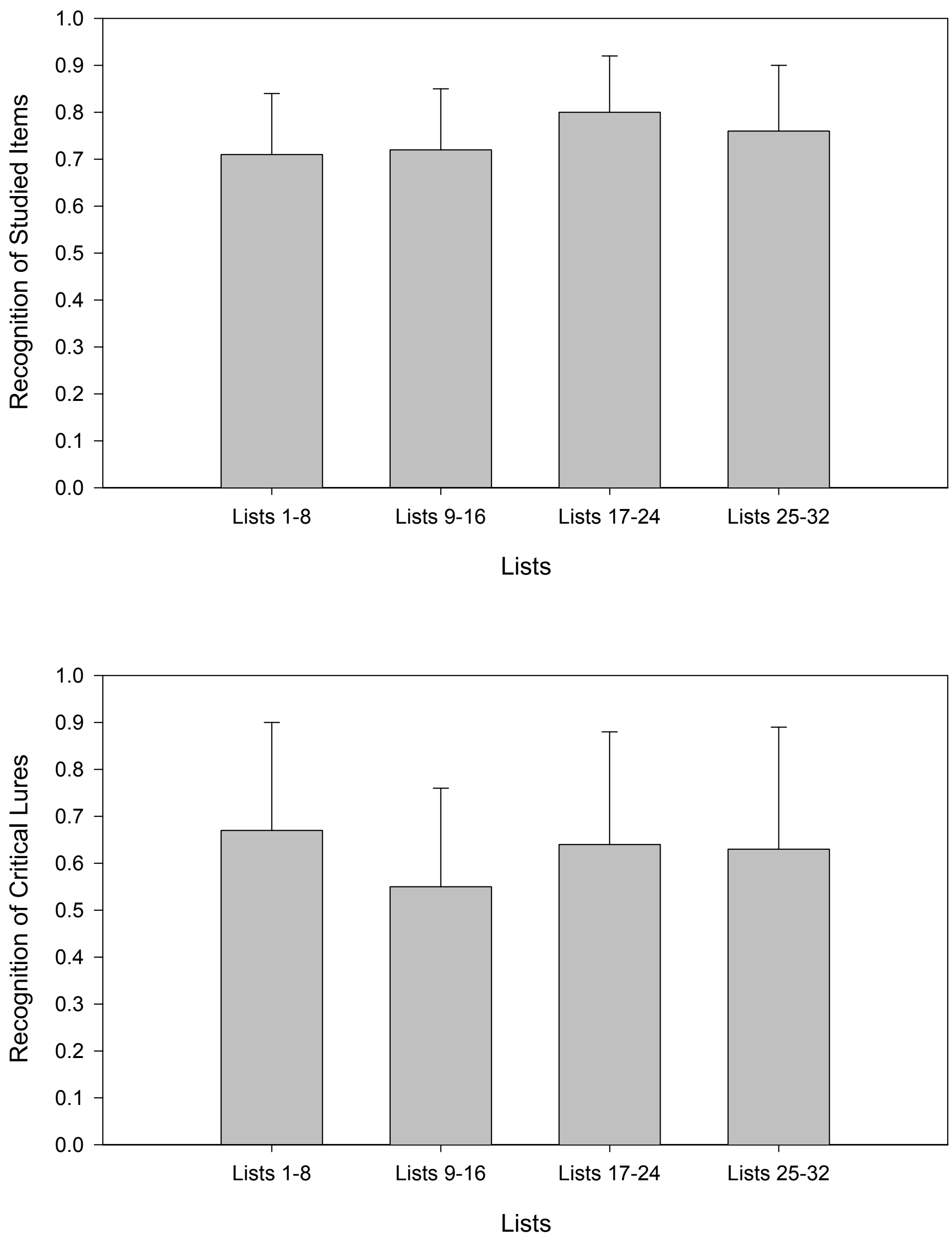


Figure 5. Actual results for the recognition of studied items (top panel) and critical lures (bottom panel).

Remember, know and guess judgments were calculated for each subject for studied items recognized as old as well as critical lures falsely recognized as old. The probability of a remember, know or guess judgment was then calculated for lists 1-8, 9-16, 17-24, and 25-32 which is presented in Table 4.

A 3 (Item: Remember, Know, Guess) x 4 (List: 1-8, 9-16, 17-24, 25-32) x 3 (Condition: High, Low, Neutral) x 2 (Test: Recall, Recognition) repeated measures ANOVA was performed. A main effect of item was significant, $F(2,110) = 15.20$, $p < .0001$, $MSE = .43$. Contrasts revealed that the remember judgments (M = .43, SD = .21) and know judgments (M= .37, SD = .21) were not significantly different ($p > .05$), but both remember and know judgments occurred at higher rates than guess judgments (M = .20, SD = .13). The test-by-item interaction was also significant, $F(2,110) = 26.15$, $p < .0001$, $MSE = .15$. Contrasts revealed that studied items were identified as old (M = .52, SD = .27) significantly more than critical lures (M = .33, SD = .23), $p < 0.05$. In addition, studied items were identified as guessed significantly less (M = .12, SD = .08) than critical lures (M = .28, SD = .22), $p < 0.05$. There was no difference between studied items and critical lures identified as known. Finally, the item-by-list interaction achieved significance, $F(6, 330) = 2.25$, $p = .0383$, $MSE = 13.03$ and is displayed in Figure 6.

Lists	Remember	Know	Guess
Studied Items			
1-8	.49 (.29)	.37 (.29)	.14 (.13)
9-16	.51 (.27)	.37 (.28)	.11 (.09)
17-24	.56 (.30)	.36 (.29)	.09 (.07)
25-32	.52 (.28)	.35 (.29)	.13 (.11)
Total	.52 (.27)	.36 (.27)	.12 (.08)
Critical Lures			
1-8	.34 (.28)	.38 (.28)	.28 (.28)
9-16	.26 (.26)	.42 (.28)	.31 (.28)
17-24	.36 (.31)	.34 (.29)	.30 (.28)
25-32	.37 (.30)	.39 (.28)	.24 (.26)
Total	.33 (.23)	.39 (.20)	.28 (.22)

Table 4. Means (Standard Errors) of Probability of Remember, Know, and Guess Judgments.

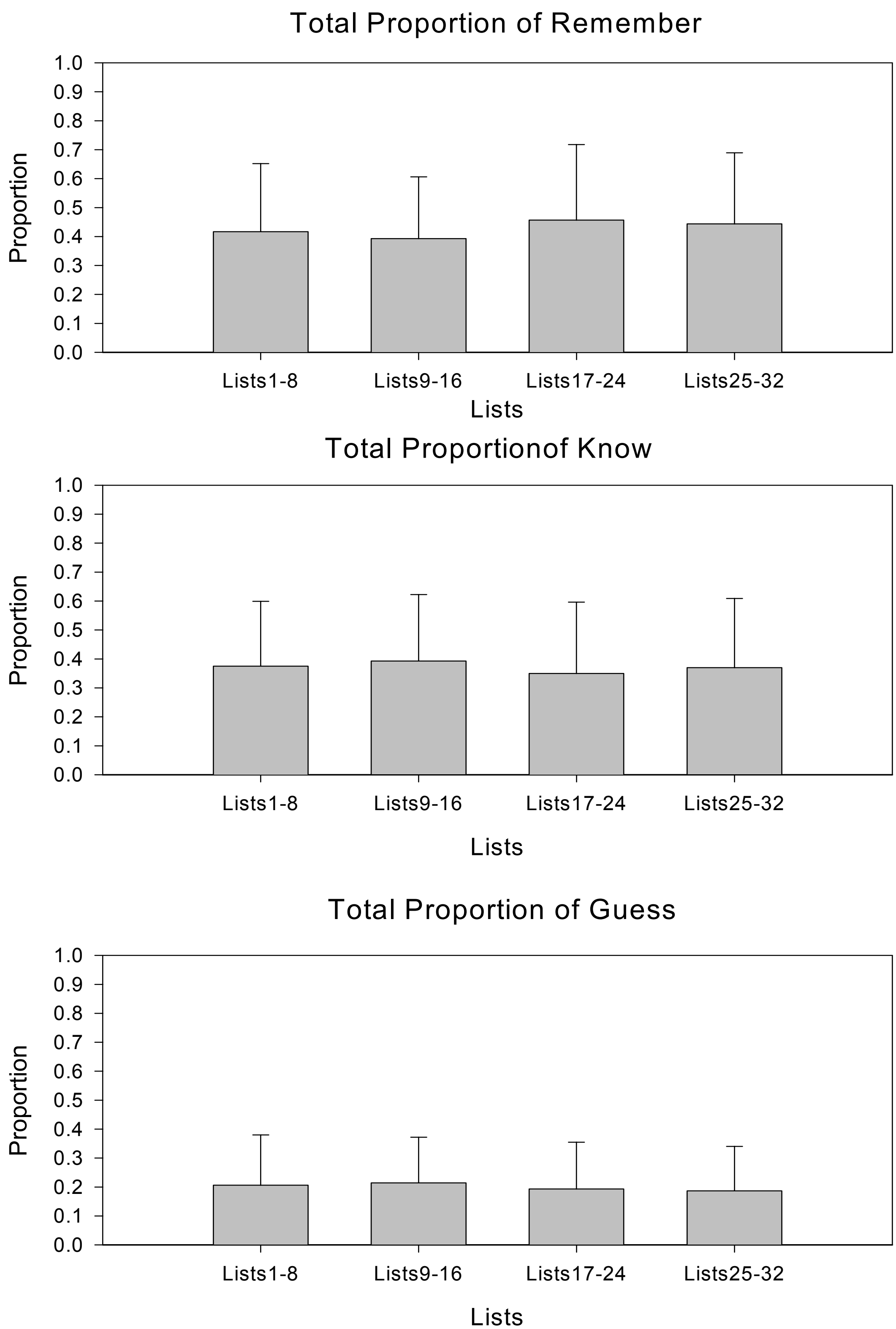


Figure 6. Proportion of remember, know, and guess responses by list.

Pre- and Post-test Memory Questionnaires

A 2 (Test: Pretest vs. Posttest) x 3 (Condition: High, Low, Neutral) repeated measures of analysis of variance (ANOVA) was performed on the overall ratings on the pre- and post-test memory questionnaire. A test by condition interaction was obtained, $\underline{F}$ (2, 57) = 13.39, $\underline{p} < .0001$, $\underline{MSE}$ = 42.55. The high-confidence group showed lower memory confidence on the pre-test questionnaire (M = 65.15, SD = 7.44) than on the post-test questionnaire (M = 71.32, SD = 7.59), $\underline{F}$ (1, 19) = 16.82, $\underline{p} < .001$, $\underline{MSE}$ = 22.67, whereas the low-confidence group showed greater memory confidence on the pre-test questionnaire (M = 67.0, SD = 10.42) than on the post-test questionnaire (M = 58.3, SD = 13.23), $\underline{F}$ (1, 19) = 14.48, $\underline{p} < .01$, $\underline{MSE}$ = 52.27. The neutral-confidence group showed no change in memory confidence from pre-to post-test ($p > .001$).

A 2 (Test: Pretest vs. Posttest) x 3 (Condition: High, Low, Neutral) repeated measures analysis of variance was performed on ratings obtained from question 1 (On a scale from 1 to 10, how would you rate your overall memory ability?) in the pre-test memory questionnaire and question 2 in the post-test memory questionnaire (Now that you've completed our memory tests, how would you rate your overall memory ability on a scale from 1 to 10). A main effect of test was obtained, $\underline{F}$ (1, 57) = 12.46, $\underline{p} < .001$, $\underline{MSE}$ = 1.27 as well as a condition by test interaction, $\underline{F}$ (2, 57) = 6.61, $\underline{p} < .01$, $\underline{MSE}$ = 1.27. Specifically, the low-confidence group showed higher confidence in general memory ability on the pre-test questionnaire (M = 6.9, SD = 1.48) than on the post-test questionnaire (M = 5.3, SD = 1.22) whereas the high and neutral confidence groups showed no differences in ratings between the two tests (both ps > 0.05).

Separate 2 (Pre- vs. Post-Test) x 3 (Condition: High, Low, Neutral) repeated measures ANOVAs were performed on the ratings obtained from each subheading of question 2 (On a scale from 1 to 10, please rate your memory ability for: keys, faces, names, phone numbers you just checked, phone numbers you use frequently, personal dates, words, remembering where your car is parked) in the pre-test memory questionnaire and question 3 (Now that you've completed our memory tests, please rate your memory ability on a scale from 1 to 10 for: keys, faces, names, phone numbers you just checked, phone numbers you use frequently, personal dates, words, remembering where your car is parked) in the post-test memory questionnaire. Table 5 displays the ANOVA results for each test.

Main effects of test were observed on the following questions: general, keys, faces, phone numbers just checked, phone numbers frequently used, words, and car. In general, ratings of memory confidence decreased for each question from pre- to post-test. Main effects of condition were also found on the following questions: phone number just checked, phone numbers used frequently, and cars. These effects were qualified by significant 2-way interactions on the following questions: general, phone numbers just checked, words, and car. For each question, memory confidence did not change between pre- and post-test for the high confidence group, whereas memory confidence significantly decreased between pre- and post-test for the low and neutral groups (all $ps < 0.05$).

	Means (SE)						2 x 3 ANOVA F-Values		
	Pre-Test			Post-Test			Cond	Test	Test x Cond
Variable	High	Low	Neut	High	Low	Neut			
General	6.35 (1.5)	6.9 (1.48)	6.75 (1.65)	6.58 (1.14)	5.3 (1.22)	5.95 (1.05)	.57	12.46*	6.61**
Keys	6.95 (1.57)	7.15 (2.08)	6.95 (1.9)	7.35 (1.44)	5.8 (5.8)	6.25 (1.29)	0.99	8.65*	7.46
Faces	8.25 (.97)	8.2 (1.85)	8.85 (1.18)	6.85 (1.79)	7.6 (1.73)	7.4 (1.39)	0.95	36.91**	2.12
Names	5.35 (1.98)	5.7 (2.32)	6.35 (2.68)	5.2 (1.44)	6.55 (2.11)	5.9 (1.33)	1.51	.11	2.47
Phone Number Just Checked	6.65 (2.16)	6.75 (2.07)	6.3 (2.23)	7.7 (1.13)	4.5 (2.26)	6.3 (1.19)	4.88*	8.18*	16.78**
Phone Number Frequent Used	9.2 (.95)	9.0 (1.3)	8.85 (1.53)	8.95 (1.0)	7.3 (2.15)	7.65 (1.35)	3.48*	28.32**	4.64
Dates	7.25 (2.31)	7.35 (7.87)	7.05 (2.39)	7.4 (1.9)	6.55 (1.79)	6.65 (1.23)	.044	1.85	1.14
Words	7.05 (1.36)	7.2 (1.91)	7.0 (2.13)	6.4 (1.27)	4.7 (2.08)	5.35 (1.53)	1.61	35.58**	3.97*
Car	8.1 (1.71)	8.75 (1.37)	8.45 (1.19)	8.55 (1.15)	5.37 (2.63)	6.5 (1.24)	4.74*	42.56**	20.00**

*$p < .005$ **$p < .001$

Table 5. Means (Standard Errors) and F-values for the Pre and Post-Test Questionnaires

Rey-Osterrieth, Intense Concentration, Digit Span, and Digit Symbol Substitution.

Z-scores were calculated separately for the Rey-Oysterrieth Test, the Intense Concentration game, the Digit Span, and Digit Symbol Substitution tests in order to compare participant's performance across each test. The resultant z-scores were submitted to a 3 (Condition: High, Low, Neutral) x 4 (Test: RO, IC, DS, DSS) repeated measures ANOVA. No significant main effects or interactions were obtained ($p > .001$). In addition, separate ANOVAs were performed on each memory test to evaluate the effect of condition on test performance. Again, no significant main effects or interactions were obtained ($p > .001$).

SUMMARY AND DISCUSSION

SUMMARY

The false memory phenomenon first revealed by Deese (1957) and further examined by Roediger and McDermott (1995) has proven to be a robust phenomenon. Many of the studies following Roediger and McDermott's (1995) revival of the study of false memories have shown that it is difficult to eliminate false memories. In fact, studies have shown that presentation rates, encoding and retrieval modality, distinctiveness, repetition, retrieval delay, social and developmental factors, and even informing participants of the false memory effect alter the magnitude, but do not eliminate the false memory effect. From this large body of research, it is clear that false memories are difficult if not impossible to abolish.

The present study sought to examine the extent to which a participant's confidence in their overall memory ability would influence the production of false memories. In addition, the present study explored the extent to which fluctuations in memory confidence influenced fluctuations in the false memory phenomenon.

Hypothesis I

It was predicted that participants would show higher rates of correct recall for studied items as compared to critical lures, as shown in previous research (e.g., Roediger & McDermott, 1995). In addition, participants were expected show higher rates of false recall of critical lures than non-studied intrusions. This hypothesis was confirmed.

Hypothesis II

It was predicted that if memory confidence affects recall, then the recall rate of studied items would change after each manipulation for both the high and low confidence groups. That is, the high confidence group should recall more studied items from lists 9-16 and 25-32 than from the lists 1-8 and 17-24, and should recall fewer studied items from lists 17-24 than any other list. The low confidence group should recall more studied items on lists 17-24 and than any other list. The low confidence group should also recall fewer studied items on lists 9-16 and 25-32 than from lists 1-8 and 17-24. The neutral group should not change in the amount of studied items recalled on any of the lists. If memory confidence affects recall, then memory confidence was anticipated to have a similar effect on the recall of critical lures.

The results obtained from this study did not confirm these predictions. Specifically, the ANOVA performed on the probability of correct recall of studied items for each set of lists did not revealed a main effect of list no main effect of condition, and no condition by list interaction.

Although no formal predictions were made, it was anticipated that false recall of critical lures would be similarly affected by the memory manipulation than studied items. In fact, the results from the ANOVA performed on the critical lures were similar to studied items. No main effect of condition or list, and no condition by list interaction were obtained.

Hypothesis III

Previous research (e.g., Roediger & McDermott, 1995) has shown higher rates of correct recognition for studied items as compared to critical lures. In addition, it was anticipated that higher rates of false recognition of critical lures would be observed compared to non-studied items. This hypothesis was confirmed.

Hypothesis IV

It was predicted that if memory confidence affects recognition, then the recognition of studied items and critical lures would change after each manipulation for both the high and low confidence groups. The high confidence group should recognize more studied items from lists 9-16 and 25-32 than from the lists 1-8 and 17-24, and should recognize fewer studied items from lists 17-24 than any other list. The low confidence should recognize more studied items from lists 17-24 and than any other list. The low confidence group should also recognize fewer studied items from lists 9-16 and 25-32 than from lists 1-8 and 17-24. The neutral group should not change in the amount of studied items recognized on any of the lists.

These predictions were not confirmed. The ANOVA performed on the recognition of studied items indicated a main effect of list, but no main effect of condition and no condition by list interaction. Similar to the recall data, there were differences in the rates of recognition as the number of lists increased, but these differences were similar for participants in the low, neutral, and high confidence conditions. Specifically, recognition probabilities remained stable from lists 1-8 to 9-16 but increased on lists 17-24 and then decrease again on lists 25-32 but not to the base level found with lists 1-8 and 9-16.

Again, although no specific predictions were made, it was expected that the recognition of critical lures would be affected by the confidence manipulation. This expectation was not borne out by the data. There was a significant decrease in the false recognition of critical lures for lists 1-8 to 9-16, while the remaining lists - lists 1-8, 17-24 and 25-32 – showed similar rates of false recognition of critical lures. No main effect of condition or condition by list interaction was observed.

Hypothesis V

It was expected that participants would make more remember judgment on studied items recognized as old, than critical lures recognized as old, consistent with results obtained by Roediger and Mc Dermott (1995). In addition, it was expected that more know responses would be made on critical lures identified as old than studied items recognized as old. The results from this study were somewhat different than the results obtained by Roediger and McDermott (1995), in that remember judgments and know judgments were not significantly different, but both remember and know judgments occurred at higher rates than guess judgments. However, studied items were identified as old significantly more than critical lures, as predicted. Studied items were also identified as guessed significantly less than critical lures. Contrary to predictions, there was no difference between studied items and critical lures identified as known.

It was also anticipated that participants in the high confidence condition would make more remember than know responses, and participants in the low confidence condition would make more know responses than remember responses. Perhaps not surprisingly, no differences were observed between the confidence groups in terms of the frequencies of remember and know responses. However, the number of guess responses was significantly less than both remember and know responses, irrespective of the confidence group.

DISCUSSION

The present experiment was designed to evaluate the extent to which the false memory phenomenon was related to metamemory, or one's confidence in one's own memory ability. Before evaluating the effects of the confidence manipulation on the false memory phenomenon, one must establish the extent to which the confidence manipulation was successful.

Participants in the high confidence condition showed an increase in scores from the pre-test to the post-test memory questionnaire, indicating that they had higher ratings of their own memory ability after the confidence manipulations. In contrast, participants in the low confidence condition showed a decrease in scores from the pre-test to the post-test memory questionnaire, indicating that they had lower ratings of their own memory ability at the end of the study. Participants in the neutral condition showed no difference between the two tests. These results provide evidence of an effective confidence manipulation.

However, analyses on the three variables manipulated on the pre-and post-test questionnaire – names, phone numbers just checked, and words – did not show consistent evidence that the confidence manipulation was successful in altering participants' meta-memory for these specific items. Specifically, there was no changed in the participants' belief to remember names, while phone numbers just checked did change in the predicted fashion, and the ratings for their ability to remember words decreased across all conditions. Thus, even though participants' perceptions in their general memory abilities seemed to be influenced by the confidence manipulation, the manipulation did not appear to consistently modify participants' beliefs about their memory abilities for specific items.

In light of evidence suggesting at least some effect of the confidence manipulation on meta-memory, the primary goal of evaluating the effects of memory confidence on false memory production can be evaluated. Although participants showed a higher probability of correctly recalling and recognizing studied items as compared to critical lures, and a higher probability of false recall and recognition of critical lures than non-critical intrusions, consistent with previous research, no significant effects of memory confidence were observed on recall or recognition of critical lures.

Several explanations may be provided for why memory confidence did not moderate recall or recognition of critical lures in the present study. The first possibility is that although general differences in memory confidence were elicited by the confidence manipulation, the manipulation may have been too weak to affect recall and recognition. It may be that participants did not entirely believe the feedback they received from the experimenter about their memory performance on each test. Alternatively, participants may have internalized the feedback from the experimenter, but getting such feedback only once was not enough to affect recall and recognition. Experiments designed to manipulate the number or amount of consistent feedback are necessary to explore this possibility.

Another possible reason why the confidence manipulation did not moderate recall or recognition of studied items or critical lures may have to do with the generalizability of the manipulation itself. That is, perhaps the confidence manipulation was successful in altering participants' meta-memory, but those alterations in memory confidence did not generalize beyond the specific memory tests used in the manipulation. For example, the WAIS-R Digit Span Subtest was used in conjunction with the variable phone numbers just checked. It is possible that participants' confidence in remembering phone numbers was affected by the manipulation, although no direct evidence was collected to determine whether this was the case. Because participants' memory abilities were not tested memory for phone numbers, but on recall and recognition of list information, the effect of the confidence manipulation did not generalize more broadly. This hypothesis could be tested in two ways: 1) test abilities specific to the tasks used in the confidence manipulation, and 2) conduct an experiment that manipulates confidence using lists (i.e., make the confidence manipulation task more like the actual memory task). The final possibility is that one's confidence in one's own memory ability is unrelated to the production of false memories. That is, participants' confidence in their memory has no impact on their memory accuracy. Therefore any manipulation done that causes the participants' confidence in their memory to be altered would not have an impact on their recall and recognition of studied items and critical lures.

Although the results revealed the manipulation was not successful in moderating the recall and recognition of critical lures, there is a possibility that the manipulation did have a more general affect. Metamemory is believed to be able to influence more general recall and recognition of items, and a comparison to the reported rates of recall and recognition by Roediger and McDermott (1995) revealed that although the rate of recall and recognition of studied items and critical lures were similar to those found by Roediger and McDermott (1995), they report higher rates of recognition (.79) of studied items and considerably higher rates of recall (.55) and recognition (.81) of critical lures than those that were found in this study (recognition of studied items = .75; recall of critical lures = .37; recognition of critical lures = .62). Such a finding may be attributed to the memory manipulation. Specifically, the post-test memory questionnaires indicated participants, overall, lost confidence in their memory. This may have manifested itself in lower rates of recall and recognition of studied items and critical lures.

An alternative explanation for this difference may be presentation modality. Roediger and McDermott (1995) presented the words aurally, while in the present study they were presented visually. Maylor and Mo (1999) found that fewer critical lures were recalled when the lists were presented visually than when they were presented aurally. In addition, they also found that correct recognition of studied items was higher with auditory presentation than with visual presentation, although they also reported higher rates of recognized critical lures with visual presentation. Future studies manipulating metamemory should evaluate if aural or visual presentation modalities affects the recall and recognition of studied items and critical lures by using a within-subjects design using both visual and aural presentations.

Future studies concerning the relationship between memory confidence and false memory research should, first, evaluate if individuals who have low confidence in their memory are more susceptible to false memory construction than individuals who have a lot of confidence in their memory. By evaluating confidence in one's own memory as a state vs. trait characteristic, one would be able to better evaluate how confidence and false memory construction are related.

Another possibility would be to concentrate on the different tasks and memory tests that could be used in attaining a manipulation of the participants' metamemory. As mentioned previously, the tests that were used in this study may not have succeeded. There are a wide variety of tests and tasks that could be used that perhaps will be more successful than the ones used in the present study. One possibility would be utilizing tests with stronger face validity with the construct being manipulated. Perhaps it would a strong manipulation would occur if the construct being manipulated would be memory in general, instead of the participants' memory for specific tasks.

In addition to using different tests to manipulate the participants' memory confidence, changing the way the memory tests are used could be another direction in research. This study attempted to manipulate the participants' confidence through common real-world scenarios. Perhaps just using their performance on the memory test, without associating it with a real-world task, would be more effective.

There has been little research outside of eyewitness testimony studies that concentrate on the proposition that one's confidence in one's own memory has an effect on performance. There are no scales to measure individuals' confidence in their memory, and there are no studies that relate this confidence in memory to performance in memory tasks. Further research in these areas should be devoted to establishing how much memory confidence or metamemory can affect memory performance. If it can be shown that metamemory does or does not moderate performance, then it would be easier to determine the extent to which memory confidence affects false memory production.

In conclusion, one can conclude there was some evidence that the memory manipulation had some affect on metamemory. However, there is evidence that the memory manipulation was not strong enough to conclude that variations in meta-memory are unrelated to the recall or recognition of studied items or critical lures. The false memory phenomenon has found to be very robust in a variety of situations, and it remains to be seen if a stronger manipulation concerning the confidence in one's own memory would have any effect on the false memory construction.

LITERATURE CITED

Anastasi, J. S., Rhodes, M. G., & Burns, M. C. (2000). Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. *American Journal of Psychology, 113,* 1-21.

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

Bechtel, W. & Abrahamson, A. (1991). <u>Connectionism and the mind: An introduction to parallel</u> processing in networks. Cambridge: Basil Blackwell.

Brainerd, C. J. & Reyna, V. F. (1990). Gist is the Grist: Fuzzy Trace Theory and the New Intuitionism. *Developmental Review, 10,* 3-47.

Brédart, S. (1999) When false memories do occur: Not thinking of the lure or remembering that it was not heard. *Memory, 8,* 123-128.

Cedrus Corporation (2000). SuperLab: Experimental Lab Software

Clark, A. (1990). Parallel distributed processing. *Microcognition*. MIT Press, pp. 83-105.

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

Craik, F. I., & Lockhart, R. S. (1973). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,*671-684.

Cutler, B. L., Penrod, S. D., & Martens. (1987b). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11*, 233-258.

Deese, J. (1957). Serial organization in the recall of disconnected items. *Psychological Reports, 3,* 577-582.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 53*, 17-22.

Deese J. & Hardman, G. W. (1954). An analysis of errors in retroactive inhibition of rote verbal learning. *American Journal of Psychology, 67*, 299-307.

Deese, J. & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology. 54*, 108-187.

Dixon, R. A., & Hultsch, D. F. (1983). Metamemory and memory for text relationships in adulthood: Cross-validation study. *Journal of Gerontology, 38*, 689-694.

Dodson, C. S. & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review, 8,* 155-161.

Drivdahl, S. B. & Zargoza, M. S. (2001). The role of perceptual elaboration and individual differences in the creation of false memories for suggested events. *Applied Cognitive Psychology, 15,* 265-281.

Gallo, D. A., Robersts, M. J. & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulleting and Review, 4*, 271-276.

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

Gallo, D. A., McDermott, K. B., Percer, J. M., & Roediger, H. L. (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 339-353.

Gardiner, J. M. & Conway, M. A. (1999). Levels of awareness and varieties of experience. In *Stratification in cognition and consciousness* ed. Challis, B. H. & Velichkaovsky, B. M. John Benjamins Publishing Company: Philadelphia

Gregory, R. L., (1968). Visual illusions. *Scientific American, 219*, 66-76.

Griffin, G. A., Glassmire, D.M., Henderson, E.A. & McCann, C. (1997). TI: Rey II: Redesigning the Rey screening test of malingering. *Journal of Clinical Psychology, 53*, 757-766.

Hall, J. F., & Kozloff, E. E. (1973). False recognition of associates of converging verses repeated words. *American Journal of Psychology, 86,* 133-139.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 528-551.

Jenkins, L., Myerson, J., Joerding, J.A., & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychology and Aging, 15*, 157-175.

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

Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The "General Acceptance" of psychological research on eyewitness testimony. *American Psychologist, 44*, 1089-1098.

Kellog, R. T. (2001) Presentation modality and mode of recall in verbal false memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 913-919.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist, 8,* 281-338.

Logie, R. H. (1995). *Visuo-Spatial Working Memory*. Hillsdale: Lawrence Erlbaum Associates.

Loring, D. W., Martin, R.C., Meador, K. J., & Lee, G.P. (1990). Psychometric construction of the Rey-Osterrieth Complex Figure: Methodological considerations and interrater reliability. *Archives of Clinical Neuropsychology, 5*, 1-14.

Maylor, E. A. & Mo, A. (1999). Effects of study-test modality in false recognition. *British Journal of Psychology, 90,* 477-493.

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

McDermott, K. B. & Roediger, H. L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under condition of explicit warnings and immediate testing. *Journal of Memory and Language, 39,* 508-520.

Migueles, M. & Garcia-Bajos, E. (1999). Recall, recognition and confidence patterns in eyewitness testimony. *Applied Cognitive Psychology, 13*, 257-268.

Miller, M. B. & Wolford, G. L. (1999). Theoretical Commentary: The role of criterion shift in false memory. *Psychological Review, 106*, 406-410.

Neil vs. Biggers. 409 U.S. 188 (1972).

Norman, K. A. & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory and Cognition, 25*, 3338-348.

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing and recollecting events that never occurred. *Journal of Memory and Language, 35,* 261-285.

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory and Cognition, 21*, 89-102.

Read, J. D. (1996). From a passing thought to a false memory in 2 minutes: Confusing real and illusory events. *Psychonomic Bulletin & Review, 3*, 105-111.

Rey, A. (1944). L'examen pschologique dans le cas d'encephalopathie traumatique. *Archives de Psychologie, 28*, 286-340.

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

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

Roediger, H. M., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8,* 365-371.

Russell, W. A. & Jenkins, J. J. (1954). The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test. (Tech. Rep. No. 11, Contract N* ONR 66216. Office of Naval Research.) Minneapolis: University of Minnesota.

Schacter, D. (1997). False recognition and the brain. *Current Directions in Psychological Science, 6,* 65-70.

Seamon, J. G, Luo, C. R., & Gallo, D. A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science, 9*, 20-26.

Schneider, C. N., Jenkins, L. A., & Ciampi, G. (in preparation). Effects of age and memory confidence on false memory construction.

Shapiro, P. N. & Penrod, S. D. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139-156.

Smith, R.E. & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5,* 710-715.

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

Steblay, Mehrkens. N. & Bothwell, R. K. (1994). Evidence for hypnotically refreshed testimony: The view from the laboratory. *Law and Human Behavior, 18*, 635-651.

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

Tulving, E. (1985). Memory and Consciousness. *Canadian Psychology, 26,* 1-13.

Tulving, E. & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 359-380.

Tun, P. A. Wingfield, A., Blanchard, L., & Rosen, M. (1996). Older adults show greater false recognition effects than young adults. Paper presented at the Cognitive Aging Conference, Atlanta, GA.

Tun, P. A. Wingfield, L., & Rosen, M. (1995). Age effects on false memory for words. Paper presented at the meeting of the Psychonomic Society, Los Angeles.

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

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30,* 128-134.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.

Wells, G.L., Lindsay, R. C. & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64,* 440-448.

Williams, S. A., Denney, N. W., & Schadler, M. (1983). Elderly adults' perception of their own cognitive development during the adult years. *International Journal of Aging and Human Development, 16*, 147-158.

Wixted, J. T. & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review, 107,* 368-376.

Appendix A

Health Questionnaire

General

1) What is your date of birth? ________________________
2) What is your sex? (circle one) Male Female
3) How many years of formal education have you completed? __________
   (12 years=high school diploma; 16 years=college graduate)

Because many of the projects we conduct are federally funded, we are required to ask about your ethnicity or racial background. You have the right to decline to answer the question that follows and if you do chose to decline, you will not be penalized in any way.

4) What is your ethnicity or race?
   ___ Caucasian (not of Hispanic descent)
   ___ Asian
   ___ American Indian
   ___ African American
   ___ Hispanic
   ___ Other

Sensory-Motor

1) Do you wear glasses or contacts? Yes No
   If so, do you wear them for: Reading
   Driving
2) When was the last time you saw your eye doctor? ______________________________
3) Do you have problems with cataracts or glaucoma Yes No
4) Have you ever had eye surgery? Yes No
   If so, describe
   ____
5) Do you have problems with tunnel vision or some other restriction of your visual field? Yes No
6) Can you see well out of both of your eyes? Yes No
7) Do you have any problems moving your fingers or arms? Yes No
8) Do you have a tremor? Yes No
   If so, how sever is the tremor? ____
   Does it affect your ability to write? Yes No

Medical

1) Are you presently under a physician's care? Yes No
   If not, when was the last time you saw a doctor? ______________________________
2) Have you been hospitalized within the last five years? Yes No
   If so, for what condition(s) ____

3) Have you ever had a heart attack or heart surgery? Yes No
If so, describe ____
4) Has a doctor ever said that you have:
a) congestive heart failure Yes No
b) diabetes Yes No
c) kidney disease Yes No
d) thyroid disease Yes No

Neurological

1) Have you ever seen a neurologist or neurosurgeon? Yes No

2) Have you ever had a sudden temporary problem with your vision, hearing, speech, memory, or ability to move? Yes No
3) Have you ever been knocked unconscious for more than 15 minutes? Yes No
4) Has a doctor ever said that you have
a) had a stroke Yes No
b) a brain tumor Yes No
c) a brain infection such as encephalitis or meningitis Yes No
d) epilepsy or seizures Yes No
e) Parkinson's Disease Yes No
f) Huntington's Disease Yes No
g) multiple sclerosis Yes No
h) Alzheimer's disease Yes No
i) any other type of head injury Yes No

Psychiatric

1) Have you ever seen a psychiatrist, psychologist, or counselor? Yes No
2) Have you ever had problems with depression? Yes No

Medications and Lifestyle

1) Approximately how many alcoholic beverages do you drink in an average week?
____
2) Have you ever had a problem with alcohol or drugs? Yes
No
3) Do you take recreational drugs (e.g., marijuana, cocaine) Yes
No
4) Are you currently taking any prescription or over-the-counter medications? Yes
No
If so, what medications are you taking?

Miscellaneous

1) Do you have any physical, emotional, or mental disabilities? Yes
No

If so, describe

2) Do you have any health problems that you think might influence your ability to participate in psychology experiments? Yes No

If so, describe

Appendix B

Memory Questionnaire

1. On a scale from 1 to 10 (with 1=poor and 10=excellent), how would you rate your overall memory ability? ____________

2. On a scale from 1 to 10 (with 1=poor and 10=excellent), please rate your memory ability for:

   a) where you put things (e.g. keys) ________
   b) faces ________
   c) names ________
   d) phone numbers you just checked ________
   e) phone numbers used frequently ________
   f) personal dates (e.g. birthdays) ________
   g) words ________
   h) remembering where you parked your car ________

3. As far as your memory ability is concerned, what do you think is your greatest strength?

   ____
   ____

   Your greatest weakness?

   ____
   ____

Appendix C

List 1 (Black) - 28
1. white
2. dark
3. cat
4. charred
5. night
6. funeral
7. color
8. grief
9. blue
10. death
11. ink
12. bottom
13. coal
14. brown
15. gray

FAS = .130

List 2 (Chair)-31
1. table
2. sit
3. legs
4. seat
5. couch
6. desk
7. recliner
8. sofa
9. wood
10. cushion
11. swivel
12. stool
13. sitting
14. rocking
15. bench

FAS = .303

List 3 (doctor)
1. nurse
2. sick
3. lawyer
4. medicine
5. health
6. hospital
7. dentist
8. physician
9. ill
10. patient
11. office
12. stethoscope
13. surgeon
14. clinic
15. cure

FAS = .245

List 4 (foot)-25
1. shoe
2. hand
3. toe
4. kick
5. sandals
6. soccer
7. yard
8. walk
9. ankle
10. arm
11. boot
12. inch
13. sock
14. smell
15. mouth

FAS = .177

List 5 (girl) - 9
1. boy
2. dolls
3. female
4. young
5. dress
6. pretty
7. hair
8. niece
9. dance
10. beautiful
11. cute
12. date
13. aunt
14. daughter
15. sister

FAS =.097

List 6 (High) - 10
1. low
2. clouds
3. up
4. tall
5. tower
6. jump
7. above
8. building
9. noon
10. cliff
11. sky
12. over
13. airplane
14. dive
15. elevate

FAS =.087

List 7 (mountain) -12

1. hill
2. valley
3. climb
4. summit
5. top
6. molehill
7. peak
8. plain
9. glacier
10. goat
11. bike
12. climber
13. range
14. steep
15. ski

FAS =.154

List 8 (needle) - 18

1. thread
2. pin
3. eye
4. sewing
5. sharp
6. point
7. prick
8. thimble
9. haystack
10. thorn
11. hurt
12. injection
13. syringe
14. cloth
15. knitting

FAS =.203

List 9 (man) -16

1. woman
2. husband
3. uncle
4. lady
5. mouse
6. male
7. father
8. strong
9. friend
10. beard
11. person
12. handsome
13. muscle
14. suit
15. old

FAS =.115

List 10 (rough)-30

1. smooth
2. bumpy
3. road
4. tough
5. sandpaper
6. jagged
7. ready
8. coarse
9. uneven
10. riders
11. rugged
12. sand
13. boards
14. ground
15. gravel

FAS =.122

List 11 (sleep) -18

1. bed
2. rest
3. awake
4. tired
5. dream
6. wake
7. snooze
8. blanket
9. doze
10. slumber
11. snore
12. nap
13. peace
14. yawn
15. drowsy

FAS =.431

List 12 (slow) -3

1. fast
2. lethargic
3. stop
4. listless
5. snail
6. cautious
7. delay
8. traffic
9. turtle
10. hesitant
11. speed
12. quick
13. sluggish
14. wait
15. molasses

FAS =.172

List 13 (soft) - 11

1. hard
2. light
3. pillow
4. plush
5. loud
6. cotton
7. fur
8. touch
9. fluffy
10. feather
11. furry
12. downy
13. kitten
14. skin
15. tender

FAS =.179

List 14 (spider) - 14

1. web
2. insect
3. bug
4. fright
5. fly
6. arachnid
7. crawl
8. tarantula
9. poison
10. bite
11. creepy
12. animal
13. ugly
14. feelers
15. small

FAS =.159

List 15 (sweet) - 4

1. sour
2. candy
3. sugar
4. bitter
5. good
6. taste
7. tooth
8. nice
9. honey
10. soda
11. chocolate
12. heart
13. cake
14. tart
15. pie

FAS =.172

List 16(thief) - 8

1. steal
2. robber
3. crook
4. burglar
5. money
6. cop
7. bad
8. rob
9. jail
10. gun
11. villain
12. crime
13. bank
14. bandit
15. criminal

FAS =.100

List 17 (anger) - 2

1. mad
2. fear
3. hate
4. rage
5. temper
6. fury
7. ire
8. wrath
9. happy
10. fight
11. hatred
12. mean
13. calm
14. emotion
15. enrage

FAS =.157

List 18 (army)

1. navy
2. soldier
3. Unites States
4. rifle
5. Air Force
6. draft
7. military
8. Marines
9. march
10. infantry
11. captain
12. war
13. uniform
14. pilot
15. combat

FAS =.135

List 19 (bread)-29

1. butter
2. food
3. eat
4. sandwich
5. rye
6. jam
7. milk
8. flour
9. jelly
10. dough
11. crust
12. slice
13. wine
14. loaf
15. toast

FAS =.200

List 20 (city) - 13

1. town
2. crowded
3. state
4. capital
5. streets
6. subway
7. country
8. New York
9. village
10. metropolis
11. big
12. Chicago
13. suburb
14. county
15. urban

FAS =.185

List 21 (flag) - 19

1. banner
2. American
3. symbol
4. stars
5. anthem
6. stripes
7. pole
8. wave
9. raised
10. national
11. checkered
12. emblem
13. sign
14. freedom
15. pendant

FAS =.109

List 22 (car) - 15

1. truck
2. bus
3. train
4. automobile
5. vehicle
6. drive
7. jeep
8. ford
9. race
10. keys
11. garage
12. highway
13. sedan
14. van
15. taxi

FAS = .347

List 23 (cold) -7

1. hot
2. snow
3. warm
4. winter
5. ice
6. wet
7. frigid
8. chilly
9. heat
10. weather
11. freeze
12. air
13. shiver
14. arctic
15. frost

FAS = .353

List 24 (fruit)

1. apple
2. vegetable
3. orange
4. kiwi
5. citrus
6. ripe
7. pear
8. banana
9. berry
10. cherry
11. basket
12. juice
13. salad
14. bowl
15. cocktail

FAS = .202

List 25 (king)-23

1. queen
2. England
3. crown
4. prince
5. George
6. dictator
7. palace
8. throne
9. chess
10. rule
11. subjects
12. monarch
13. royal
14. leader
15. reign

FAS = .230

List 26 (music) - 6

1. note
2. sound
3. piano
4. sing
5. radio
6. band
7. melody
8. horn
9. concert
10. instrument
11. symphony
12. jazz
13. orchestra
14. art
15. rhythm

FAS = .227

List 27 (river)

1. water
2. stream
3. lake
4. Mississippi
5. boat
6. tide
7. swim
8. flow
9. run
10. barge
11. creek
12. brook
13. fish
14. bridge
15. winding

FAS = .147

List 28 (lion) - 5

1. tiger
2. circus
3. jungle
4. tamer
5. den
6. cub
7. Africa
8. mane
9. cage
10. feline
11. roar
12. fierce
13. bears
14. hunt
15. pride

FAS = .136

List 29 (pen) -22

1. pencil
2. write
3. fountain
4. leak
5. quill
6. felt
7. bic
8. scribble
9. crayon
10. cross
11. tip
12. marker
13. red
14. cap
15. letter

FAS =.135

List 30 (rubber)-24

1. elastic
2. bounce
3. gloves
4. tire
5. ball
6. eraser
7. springy
8. foam
9. galoshes
10. soles
11. latex
12. glue
13. flexible
14. resilient
15. stretch

FAS = .033

List 31(smell)-21

1. nose
2. breathe
3. sniff
4. aroma
5. hear
6. see
7. nostril
8. whiff
9. scent
10. reek
11. stench
12. fragrance
13. perfume
14. salts
15. rose

FAS = .290

List 32 (trash)- 27

1. garbage
2. waste
3. can
4. refuse
5. sewage
6. bag
7. junk
8. rubbish
9. sweep
10. scraps
11. pile
12. dump
13. landfill
14. debris
15. litter

FAS = .140

List 33 (cup) - 20

1. mug
2. saucer
3. tea
4. measuring
5. coaster
6. lid
7. handle
8. coffee
9. straw
10. goblet
11. soup
12. stein
13. drink
14. plastic
15. sip

FAS =.109

List 34 (smoke)-25

1. cigarette
2. puff
3. blaze
4. billows
5. pollution
6. ashes
7. cigar
8. chimney
9. fire
10. tobacco
11. stink
12. pipe
13. lungs
14. flames
15. stain

FAS = .167

List 35 (shirt) - 1

1. blouse
2. sleeves
3. pants
4. tie
5. button
6. shorts
7. iron
8. polo
9. collar
10. vest
11. pocket
12. jersey
13. belt
14. linen
15. cuffs

FAS =.186

List 36 (window)-32

1. door
2. glass
3. pane
4. shade
5. ledge
6. sill
7. house
8. open
9. curtain
10. frame
11. view
12. breeze
13. sash
14. screen
15. shutter

FAS = .184

List 37 (short)
1. tall
2. long
3. fat
4. small
5. hair
6. stout
7. stocky
8. shirt
9. stubby
10. little
11. dwarf
12. tiny
13. midget
14. curt
15. brief

FAS =?

List 38 (wish)
1. want
2. desire
3. hope
4. well
5. think
6. star
7. bore
8. ring
9. wash
10. thought
11. money
12. true
13. Christmas
14. birthday
15. fairy

FAS = .012

List 39 (whiskey)
1. drink
2. drunk
3. beer
4. liquor
5. gin
6. bottles
7. alcohol
8. scotch
9. wine
10. rum
11. bourbon
12. drunkard
13. brandy
14. booze
15. Irish

FAS = .022

List 40 (whistle)
1. stop
2. train
3. noise
4. sing
5. blow
6. loud
7. shrill
8. toot
9. tweet
10. siren
11. tin
12. pucker
13. lips
14. bell
15. holler

FAS = .005

List 41 (carpet)
1. rug
2. floor
3. red
4. sweeper
5. tack
6. bagger
7. room
8. house
9. thick
10. sweep
11. wool
12. plush
13. weave
14. Persian
15. vacuum

FAS =.037

List 42 (square)
1. round
2. circle
3. block
4. box
5. root
6. triangle
7. cube
8. rectangle
9. time
10. geometry
11. corner
12. oblong
13. mile
14. angle
15. dumb

FAS =?

List 43 (Earth)
1. round
2. dirt
3. ground
4. moon
5. planet
6. world
7. land
8. soil
9. sun
10. heaven
11. Mars
12. Universe
13. star
14. sphere
15. quake

FAS = ?

List 44 (eagle)
1. bird
2. fly
3. scout
4. wings
5. nest
6. bald
7. hawk
8. beak
9. soar
10. feathers
11. flight
12. coin
13. clan
14. eye
15. falcon

FAS = ?

List 45 (moon)
1. stars
2. sun
3. light
4. shine
5. beam
6. bright
7. glow
8. romance
9. full
10. crescent
11. tide
12. face
13. telescope
14. luna
15. astronomy

FAS = ?

List 46 (cabbage)
1. lettuce
2. salad
3. food
4. eat
5. green
6. garden
7. leaf
8. sauerkraut
9. slaw
10. patch
11. plant
12. soup
13. ham
14. head
15. stew

FAS =?

List 47 (Bible)
1. God
2. book
3. church
4. religion
5. holy
6. read
7. Jesus
8. Christ
9. Moses
10. faith
11. prayer
12. testament
13. word
14. lord
15. psalms

FAS =?

List 48 (hammer)
1. nail
2. saw
3. hit
4. pound
5. tool
6. carpenter
7. iron
8. chisel
9. build
10. beat
11. throw
12. sledge
13. repair
14. thumb
15. heavy

FAS =?

List 1 -8:
1. shirt
2. anger
3. slow
4. sweet
5. lion
6. music
7.cold
8.thief

List 9-16
9.girl
10. anger
11. soft
12. mountain
13. city
14. spider
15. car
16. man

List 17-24
17. needle
18. sleep
19.flag
20. cup
21. smell
22. pen
23. king
24. rubber

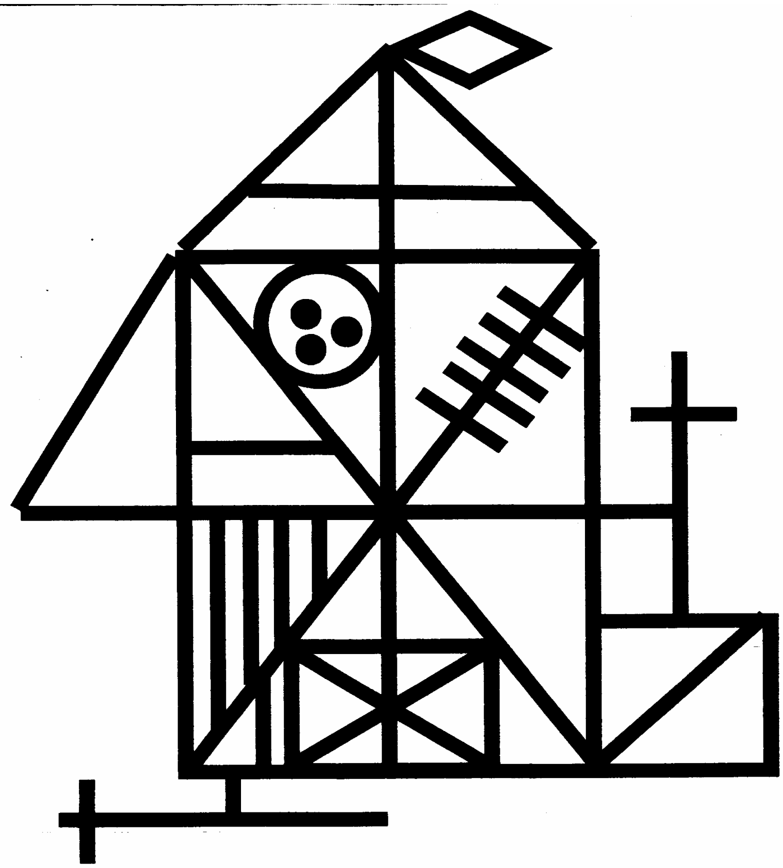
List 25-32
25. foot
26. smoke
27. trash
28. black
29. bread
30. rough
31. chair
32. window

Lists not presented:

river
army
fruit
short
wish
whiskey
doctor
hammer

whistle
carpet
square
earth
eagle
moon
cabbage
bible

Appendix D



Appendix E

DISCONTINUE RULE
Digits Forward & Backward:
Score of 0 on both trials of any item
For both Digits Forward & Backward, administer both trials of each item even if Trial 1 is passed. Administer Digits Backward even if examinee scores 0 on Digits Forward.

SCORING RULE
Each Trial: 0 or 1 pt. for each response
Item score = Trial 1 + Trial 2

Digits Forward

Item	Trial	Item/Response	Trial Score	Item Score (0, 1, or 2)
1.	1	1-7		
	2	6-3		
2.	1	5-8-2		
	2	6-9-4		
3.	1	6-4-3-9		
	2	7-2-8-6		
4.	1	4-2-7-3-1		
	2	7-5-8-3-6		
5.	1	6-1-9-4-7-3		
	2	3-9-2-4-8-7		
6.	1	5-9-1-7-4-2-8		
	2	4-1-7-9-3-8-6		
7.	1	5-8-1-9-2-6-4-7		
	2	3-8-2-9-5-1-7-4		
8.	1	2-7-5-8-6-2-5-8-4		
	2	7-1-3-9-4-2-5-6-8		

Digits Forward Total Score (Maximum = 16)

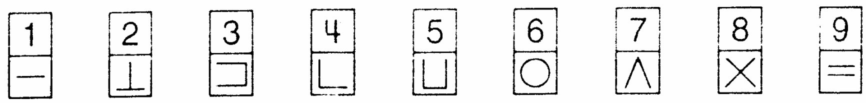
Digits Backward

Item	Trial	Item/Response	Trial Score	Item Score (0, 1, or 2)
1.	1	2-4		
	2	5-7		
2.	1	6-2-9		
	2	4-1-5		
3.	1	3-2-7-9		
	2	4-9-6-8		
4.	1	1-5-2-8-6		
	2	6-1-8-4-3		
5.	1	5-3-9-4-1-8		
	2	7-2-4-8-5-6		
6.	1	8-1-2-9-3-6-5		
	2	4-7-3-9-1-2-8		
7.	1	9-4-3-7-6-2-5-8		
	2	7-2-8-1-9-6-5-3		

Digits Backward Total Score (Maximum = 14)

Forward + Backward = (Maximum = 30)

Digit Symbol—Coding



Sample Items

2	1	3	7	2	4	8	2	1	3	2	1	4	2	3	5	2	3	1	4

5	6	3	1	4	1	5	4	2	7	6	3	5	7	2	8	5	4	6	3

7	2	8	1	9	5	8	4	7	3	6	2	5	1	9	2	8	3	7	4

6	5	9	4	8	3	7	2	6	1	5	4	6	3	7	9	2	8	1	7

9	4	6	8	5	9	7	1	8	5	2	9	4	8	6	3	7	9	8	6

2	7	3	6	5	1	9	8	4	5	7	3	1	4	8	7	9	1	4	5

7	1	8	2	9	3	6	7	2	8	5	2	3	1	4	8	4	2	7	6

Appendix G

Recognition Test

Instructions: For every word, decide if it has been presented (old) or not (new). If the word is old make the distinction if you remember the word, you know or guess it. Remembering the words indicates you remember something specific about the word (for example where it was in the list or remember picturing the word). Knowing the word means that you can't remember any specifics, but you know the word was presented, and guessing means you are guessing the word was on one of the lists.

Word	(circle one)		**Remember (R)/ Know (K) or Guess (G)**
1. anger	old	new	______________
2. tender	old	new	______________
3. old	old	new	______________
4. pie	old	new	______________
5. siren	old	new	______________
6. bed	old	new	______________
7. elastic	old	new	______________
8. patch	old	new	______________
9. banana	old	new	______________
10. man	old	new	______________
11. note	old	new	______________
12. whiskey	old	new	______________
13. infantry	old	new	______________
14. door	old	new	______________
15. whistle	old	new	______________
16. barge	old	new	______________
17. physician	old	new	______________
18. cabbage	old	new	______________
19. animal	old	new	______________
20. stain	old	new	______________
21. shirt	old	new	______________
22. table	old	new	______________
23. flow	old	new	______________
24. cold	old	new	______________
25. eagle	old	new	______________
26. rug	old	new	______________
27. rough	old	new	______________
28. beak	old	new	______________

29. nose	old	new	________
30. God	old	new	________
31. shutter	old	new	________
32. Chicago	old	new	________
33. sip	old	new	________
34. scotch	old	new	________
35. litter	old	new	________
36. soft	old	new	________
37. doctor	old	new	________
38. slow	old	new	________
39. rose	old	new	________
40. enrage	old	new	________
41. monarch	old	new	________
42. pendant	old	new	________
43. thread	old	new	________
44. pipe	old	new	________
45. nap	old	new	________
46. rum	old	new	________
47. downy	old	new	________
48. bread	old	new	________
49. soil	old	new	________
50. city	old	new	________
51. letter	old	new	________
52. want	old	new	________
53. town	old	new	________
54. banner	old	new	________
55. quick	old	new	________
56. rubber	old	new	________
57. smell	old	new	________
58. climber	old	new	________
59. highway	old	new	________
60. stool	old	new	________
61. king	old	new	________
62. heaven	old	new	________
63. belt	old	new	________
64. cherry	old	new	________
65. breeze	old	new	________
66. sauerkraut	old	new	________
67. toast	old	new	________
68. foot	old	new	________
69. crescent	old	new	________
70. garbage	old	new	________
71. blouse	old	new	________

72. apple	old	new	______
73. butter	old	new	______
74. flag	old	new	______
75. river	old	new	______
76. chisel	old	new	______
77. lettuce	old	new	______
78. window	old	new	______
79. hot	old	new	______
80. pencil	old	new	______
81. nail	old	new	______
82. mean	old	new	______
83. sleeves	old	new	______
84. Christ	old	new	______
85. molasses	old	new	______
86. fast	old	new	______
87. moon	old	new	______
88. cup	old	new	______
89. romance	old	new	______
90. wish	old	new	______
91. glue	old	new	______
92. stars	old	new	______
93. Marines	old	new	______
94. heart	old	new	______
95. pride	old	new	______
96. thief	old	new	______
97. army	old	new	______
98. low	old	new	______
99. trash	old	new	______
100. hammer	old	new	______
101. rectangle	old	new	______
102. sweet	old	new	______
103. nurse	old	new	______
104. high	old	new	______
105. hard	old	new	______
106. marker	old	new	______
107. chair	old	new	______
108. water	old	new	______
109. sister	old	new	______
110. spider	old	new	______
111. stein	old	new	______
112. criminal	old	new	______
113. slice	old	new	______
114. stop	old	new	______

115. drink	old	new	______
116. round	old	new	______
117. little	old	new	______
118. carpet	old	new	______
119. crime	old	new	______
120. knitting	old	new	______
121. date	old	new	______
122. patient	old	new	______
123. black	old	new	______
124. girl	old	new	______
125. tiger	old	new	______
126. air	old	new	______
127. boy	old	new	______
128. music	old	new	______
129. handsome	old	new	______
130. mountain	old	new	______
131. jazz	old	new	______
132. emblem	old	new	______
133. short	old	new	______
134. fruit	old	new	______
135. toot	old	new	______
136. fierce	old	new	______
137. taxi	old	new	______
138. dump	old	new	______
139. shoe	old	new	______
140. ski	old	new	______
141. rhythm	old	new	______
142. stretch	old	new	______
143. injection	old	new	______
144. sleep	old	new	______
145. sweep	old	new	______
146. Bible	old	new	______
147. truck	old	new	______
148. bottom	old	new	______
149. woman	old	new	______
150. gray	old	new	______
151. small	old	new	______
152. beat	old	new	______
153. bird	old	new	______
154. hill	old	new	______
155. over	old	new	______
156. smooth	old	new	______
157. dirt	old	new	______

158. mad	old	new	________
159. navy	old	new	________
160. needle	old	new	________
161. lion	old	new	________
162. geometry	old	new	________
163. tall	old	new	________
164. car	old	new	________
165. elevate	old	new	________
166. white	old	new	________
167. ring	old	new	________
168. sour	old	new	________
169. faith	old	new	________
170. drowsy	old	new	________
171. web	old	new	________
172. steal	old	new	________
173. feathers	old	new	________
174. pen	old	new	________
175. frost	old	new	________
176. mug	old	new	________
177. gravel	old	new	________
178. inch	old	new	________
179. shirt	old	new	________
180. queen	old	new	________
181. urban	old	new	________
182. square	old	new	________
183. fragrance	old	new	________
184. house	old	new	________
185. bench	old	new	________
186. thought	old	new	________
187. earth	old	new	________
188. sand	old	new	________
189. cigarette	old	new	________
190. reign	old	new	________
191. smoke	old	new	________
192. mouth	old	new	________

Appendix H

Memory Questionnaire

1. On a scale from 1 to 10 (with 1=poor and 10=excellent), how well do you think you performed on our memory tests overall? ____________

2. Now that you've completed our memory tests, how would you rate your overall memory ability on a scale from 1 to 10 (with 1=poor and 10=excellent)? ____________

3. Now that you've completed our memory tests, please rate your memory ability on a scale from 1 to 10 (with 1=poor and 10=excellent) for:

   a) where you put things (e.g. keys) ________
   b) faces ________
   c) names ________
   d) phone numbers you just checked ________
   e) phone numbers used frequently ________
   f) personal dates (e.g. birthdays) ________
   g) words ________
   h) remembering where you parked your car ________

4. Now that you've completed our memory tests, please consider your overall memory ability. What do you think is your greatest strength?
   ____
   ____

   Your greatest weakness?
   ____
   ____

5. Below, describe what do you think this experiment is about.