

# Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings

JEFFREY S. ANASTASI, MATTHEW G. RHODES, AND  
MATTHEW C. BURNS  
Francis Marion University

Previous research has demonstrated that the false memory effect is robust and that false memories are essentially indistinguishable from memories for events that actually occurred. The current study used several techniques intended to eliminate false memories (source monitoring decisions, confidence ratings, re-member/know judgments, and explicit warnings). A robust false memory effect was found in each experiment. However, participants were able to differentiate false memories and actual memories when using specific phenomenological tasks. The current findings provide insight into basic human memory processes.

Recently there has been considerable interest in the idea of false memories or memories for events that never actually occurred. Although there has been a recent upsurge in research investigating memory illusions, the phenomenon is not new. In fact, the literature contains numerous examples of both anecdotal and experimental evidence supporting the existence of false memories (Anastasi, Burns, & Golden, 1996; Ceci, Huffman, & Smith, 1994; Deese, 1959; Gallo, Roberts, & Seamon, 1997; Lindsay & Read, 1994; Loftus, 1993; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996; Payne, Neuschatz, Lampinen, & Lynn, 1997; Piaget, 1951; Pynoos & Nader, 1989; Rhodes & Anastasi, in press; Roediger & McDermott, 1995; Schooler, 1994; Seamon, Luo, & Gallo, 1998). Obviously, an experimental study that directly investigated the validity of false memories in the context of a very high stress episode would be unethical. In addition, if one were to investigate real-world situations, then one would have to wait for these types of situations to occur naturally, rendering the investigation of false memories extremely time-consuming and unpredictable. Therefore, a method for producing false memories that can be carried out in the laboratory is required so that we may better understand the causes of the false memory phenomenon.

Deese (1959) demonstrated a method that reliably produces false memories and can be used to investigate them. In Deese's study, participants were presented with a list of words that were associates of a single, nonpresented item. For example, participants may be presented with a list of words that are all related to the word *sleep*, such as *bed*, *dream*, *snore*, and *awake*, but did not contain the critical lure *sleep*. Results indicate that participants remembered many of the sleep words that were presented and also remembered *sleep* as being one of the presented items. Deese provides a technique for reliably producing false memories for items that never appeared and allows researchers to accurately predict the identity of the nonlist intrusions (i.e., the critical lures). Therefore, the Deese paradigm demonstrates that memory is not a verbatim reproduction but rather an imperfect representation.

Roediger and McDermott (1995) extended Deese's (1959) research by replicating his results using recall tests and extended his findings to recognition tests. Roediger and McDermott also found that participants maintained a high degree of confidence for the nonpresented critical lures. These confidence ratings indicate that not only did participants remember the critical lures, but also that their phenomenological experience was similar to their memories for words that were actually presented. Furthermore, Roediger and McDermott provided additional evidence indicating a similar phenomenological representation for memory illusions and actual memories by using remember/know judgments. Remember judgments are made when participants are able to reexperience the original episode, whereas know judgments are made when participants are very confident that the item was presented earlier but cannot remember the specific episode (Thlvig, 1985). Results revealed that participants made as many remember judgments for the critical lures as they did for the list items. In other words, participants seemed to remember episodic details for items that were never actually presented. Roediger and McDermott's findings thus support the idea that not only do participants remember the nonpresented items but these items also seem to have a memory quality similar to items that were actually presented.

Based on these results, Roediger and McDermott concluded that the false memory effect observed in their studies and Deese's (1959) experiment involve a memory process such that the "memory" for the nonpresented items is indistinguishable from memories for words that were actually presented. More recently, Schacter (1996) used neuroimaging techniques and demonstrated that illusory memories seem to activate many of the same areas in the brain as actual memories, indicating that false memories are indistinguishable from memories for actual events. Although many earlier studies have not been able to distinguish be-

tween true and false memories, some recent research has shown differences between true and false memories. Specifically, Mather, Henkel, and Johnson (1997) used a memory characteristics questionnaire to assess phenomenological memory quality. Their findings indicate that although there was no statistical difference in confidence ratings between true and false memories, participants' evaluations of false memories contained less auditory detail than actual memories. Norman and Schacter (1997) also found phenomenological differences between true and false memories. Their primary differences were for sound, list position, and reaction judgments. Furthermore, these findings are consistent with those of Schacter, Reiman, Curran, Yun, Bandy, McDermott, and Roediger (1996), who found that veridical recognition was accompanied by temporoparietal brain activation, which has been linked to phonological processing. Notwithstanding, the extant research is mixed with regard to identifying the characteristics that distinguish between illusory and veridical memories.

The current study used several different memory tasks to provide participants with the opportunity to examine the phenomenological nature of their memories. Specifically, the techniques used were source monitoring decisions (Experiment 1), confidence ratings (Experiment 2), explicit warnings (Experiments 2 and 3), and remember/know judgments (Experiment 3). The phenomenological characteristics of both the false memories and memories for actual events, rather than just overall recognition accuracy scores, were investigated. As previous research has shown, participants' recognition accuracy may be somewhat equivalent across item types (i.e., list vs. critical lures) whereas participants' memory for the critical lures may be less distinct or have fewer details (Mather et al., 1997; Norman & Schacter, 1997). Moreover, it has been demonstrated that a reliance on recognition accuracy alone may not provide enough information about the quality of participants' memory for the original items or events. In the current study, participants' recognition memory was evaluated by using various testing techniques to determine the quality of the original memory. Several studies have established that confidence and recognition judgments are not correlated and may reflect different aspects of memory (Loftus, Donders, Hoffman, & Schooler, 1989; Smith, Kassir, & Ellsworth, 1989; Zaragoza & Koshmider, 1989). Therefore, recognition tests may measure an overall familiarity with the items whereas source monitoring judgments, confidence ratings, and remember/know judgments measure other phenomenological characteristics of memory (Anastasi et al., 1996; Payne et al., 1996). If participants' memory for the critical lures and list items is indistinguishable then one would expect no difference in these quality judgments. However, differences in these quality judgments

would seem to indicate real differences between participants' memory for the critical lures and the list items.

#### EXPERIMENT 1

Eyewitness memory research provides an excellent example of methods that can be applied to the false memory literature. Researchers in this area have faced dilemmas similar to those that false memory researchers are currently investigating. One main concern expressed in the eyewitness literature is whether witnesses are able to distinguish between information that was actually seen and information that was merely suggested (Christiansen & Ochalek, 1983; Greene, Flynn, & Loftus, 1982; Lindsay & Johnson, 1989).

Several studies have demonstrated that the postevent suggestibility effect is fairly robust (Belli, Windschitl, McCarthy, & Winfrey, 1992; Chandler, 1989; Loftus, 1975, 1979; Loftus & Palmer, 1974; Payne, Toglia, & Anastasi, 1994; Toglia, Ross, Ceci, & Hembrooke, 1992). Additionally, Lindsay and Johnson (1989) have investigated the misleading postevent suggestibility effect in eyewitness memory using source monitoring judgments. Their study demonstrated that participants seem to accept the misleading information unless they are specifically instructed to determine the source of the information. According to Lindsay and Johnson, participants typically do not use source monitoring processing unless they are specifically required to do so. However, when participants are asked to make source monitoring decisions, the strength of the misinformation effect is diminished or even disappears. The source monitoring task thus allows participants to be more critical of the source of the information rather than simply accept information that the experimenter suggests as veridical.

A similar type of finding may occur within the false memory literature. Memories for items that never actually occurred may become distinguishable from memory for items that were actually presented when participants are required to engage in source monitoring processing. Payne et al. (1996) investigated the ability of participants to use source monitoring processing using a false memory paradigm. Their goal was to determine whether participants could distinguish between items that were actually presented and false memories. Payne et al. provided two different sources of information by having a list of words presented using both a male and female speaker. After the list presentation, participants were then given three successive recall tests. Following the third recall test, participants were asked to make a source attribution for each recalled item. Their results indicate that participants attributed list items

and critical lures to a source 94% and 87% of the time, respectively. In other words, participants were very likely to attribute recalled items to a source whether these items were presented on the original word list or not. However, there are three primary concerns with their method.

Payne et al. (1996) may be biasing participants to attribute recalled items to a source because these items may seem very familiar. First, Payne et al. used a recall procedure that may bias participants' source judgments. If the item was recalled, then participants would think that the item was presented and would be more likely to attribute this item to a particular source. A recognition test contains both items that were on the list and items that were not on the list so this bias is eliminated or weakened. Second, participants in the Payne et al. study made their source judgments *after* they had retrieved the list items. This may have made them more likely to attribute these items to a particular source because these items already seemed familiar. Finally, they allowed participants to use source monitoring after the third recall test in their hypernesia paradigm. Again, any item that was recalled on previous lists may have been more familiar because of prior retrievals. The current study required participants to engage in source monitoring judgments as the only type of retrieval, thus minimizing any chances of bias.

Overall, the purpose of Experiment 1 was to test the robustness of the Payne et al. (1996) findings using a recognition test. It was expected that source monitoring judgments would allow participants to distinguish between false memories and memories of actual events by requiring participants to critically evaluate the source of the information. Moreover, the current experiment required participants to engage in source monitoring as the only retrieval task.

#### METHOD

##### Participants

Twenty participants enrolled in an introductory psychology course at Francis Marion University received partial credit for a library research requirement or as class extra credit. Participants were tested in groups of two to six.

##### Materials

The materials used in the current experiment were taken from the Nelson, McEvoy, and Schreiber (1994) association word lists. Four 14-item sublists were used in the current experiment. The sublists were blocked so that all of the items from each sublist were followed immediately by all of the items from the next sublist, until all four sublists were presented. The words within each sublist were in a random order. A pilot study determined that the Nelson et al. word lists produced a false memory effect of a similar magnitude as the Deese (1959)

word lists. Participants were very likely to recognize both the list items and the critical lures (.89 and .92, respectively) and very unlikely to recognize the non-list distractors (.07). Results indicated that there was no significant difference between recognition of list items and critical lures,  $t(11) = -0.73$ ,  $p = .48$ . However, the nonlist distractors were significantly different from both the critical lures,  $t(11) = 14.10$ , and the list items,  $t(11) = 19.28$ .

The word fragment completion task was taken from Tulving, Schacter, and Stark (1982) and consisted of 104 fragments, each of which was missing three letters. This task served as an unrelated filler that would eliminate rehearsal of the list items.

The 24-item source monitoring test was composed of 12 presented items and 12 nonpresented items. Of the test items, 12 items were taken from the original lists (three items randomly selected from each of the four presented lists), 6 items were randomly selected from two nonpresented lists, 2 items were the critical lures from the nonpresented lists, and the final 4 items were the critical lures from the four presented lists. For the *flower* list, an example of a list item would be *blossom*, an example of a nonlist item would be *beach* (a nonrelated item from one of the nonpresented lists), and the critical lure would be *flower*. The critical lures were items that were never actually presented on the original list presentation but each of the list items was associated with these critical lures. The order of test items was random. Appendix A provides each of the word lists used in the current study and indicates the items used on the recognition tests.

#### Procedure

Participants were shown a videotape on which a Caucasian man and an African-American woman read the list items at a rate of approximately 1.5 s per word. The tape showed each speaker's shoulders and head so that participants could see the speakers' faces as they read each word. Half of the words from each list were spoken by the man and half were spoken by the woman. This was counterbalanced so that each word was spoken by the man half of the time and by the woman half of the time. Participants viewed the words from seats that were located 3–5 feet away from the television screen.

After the list presentation, participants were given the word fragment completion task, which served as an unrelated filler task that would eliminate rehearsal of the list items. Participants were instructed to add letters to complete as many of the word fragments as possible in the allotted 8 min.

Once participants completed the word fragment task, they were presented with a 24-item source monitoring test. Participants were given instructions on how to make the source judgments. At the top of the sheet was a 4-point scale. If participants felt that a particular word was spoken by the man, they were instructed to write an *M* in the space next to the word. If they felt that the word was spoken by the woman, they were to write an *F* in the space next to the word. If they felt that the particular word was presented but were unsure who said it, they were asked to write a *U* in the space next to the word. If participants felt that a particular word was not actually presented on the original list, they were asked to write *NP* in the blank provided.

## RESULTS AND DISCUSSION

The first analyses were conducted to determine the probability of calling an item "old." Although participants did not specifically make a yes/no or old/new judgment, one can determine this probability by using the source judgments. On the source monitoring test, participants were asked to determine the source of the item or indicate whether the item was not actually presented. Therefore, any rating indicating that the item was presented (i.e., man [M], woman [F], unsure [U]) would be considered an "old" response and any rating indicating that the item was not actually presented (i.e., *NP*) would correspond to a "new" response.

The probability of calling list items and critical lures "old" was .92 and .93, respectively, and nonlist distractors were called "old" less often (.37). A one-way ANOVA demonstrated an effect of item type,  $F(2, 38) = 44.02$ ,  $MSE = 0.05$ ,  $p < .01$  (an alpha level of .01 was used for all statistical tests unless otherwise noted). Follow-up *t* tests indicated that the difference between the probability of calling a list item and a critical lure "old" was not significant,  $t(19) = -0.19$ ,  $p = .86$ . However, the difference between the list items and nonlist distractors was significant,  $t(19) = 7.13$ , as was the difference between the critical lures and the nonlist distractors,  $t(19) = 6.49$ . Therefore, the list items and critical lures seem to be indistinguishable when looking at the recognition data. Even though participants were required to use source monitoring processing, they continued to demonstrate an inability to distinguish between the critical lures and list items. To look at the direct effects of source monitoring processing, one must look at the specific source monitoring data.

We must first determine whether participants are able to assign the list items to the correct source (e.g., male or female speaker). Although this was a difficult task, participants attributed the list items to their proper source with 60% accuracy even after performing the 8-min filler task. Of course, the main question that must be answered is whether participants were able to distinguish the critical lures from the list items when making source judgments.

A one-way ANOVA was conducted to determine whether participants attributed the critical lures to a source similar to the list items or whether they were able to distinguish between the critical lures and list items. The probability of attributing previously recognized list items, critical lures, and nonlist distractors to a source was .83, .78, and .34, respectively. Results revealed a main effect of item type,  $F(2, 38) = 24.09$ ,  $MSE = 0.06$ . Follow-up *t* tests demonstrated that, as expected, the list items were attributed to a source more often than the nonlist distractors,  $t(19) = 6.20$ . More importantly, the critical lures were attributed to a source

more often than the nonlist distractors,  $t(19) = 4.66$ , but participants were just as likely to attribute the critical lures and list items to a source,  $t(19) = 0.89$ ,  $p = .39$ . The current findings support the results of Payne et al. (1996) because participants were very likely to attribute critical lures to a particular source even though the critical lures were not actually presented. Even though Payne et al. used a different method, the results of the current study substantiate their findings.

Requiring participants to specifically engage in source monitoring did not have an effect on their performance. Based on the research conducted by Lindsay and Johnson (1989), it was expected that participants would be able to distinguish between the critical lures and actual list items when asked to perform source monitoring processing. This is not what was found. Participants in Experiment 1 not only remembered the critical lures as previously presented items, but also assigned them to a particular source. It is surprising that participants continued to assign a source to items that were never actually presented, even when using a recognition test. These results demonstrate that the false memory effect is indeed robust and show the clarity of these false memories. In addition, false memories seem to be indistinguishable from memories of items that were actually presented.

It was assumed that the source monitoring instructions in Experiment 1 would provide an implicit warning to participants that some of the items were not actually presented on the original list, but either participants did not heed this warning or they were not able to distinguish between critical lures and list items. Experiment 2 explored the possibility that a more explicit warning may be required for participants to distinguish between actual and false memories.

## EXPERIMENT 2

Several researchers have used an explicit warning to reduce the effects of misleading postevent information in eyewitness studies (Anastasi, 1992; Christiansen & Ochalek, 1983; Greene et al., 1982). In misinformation eyewitness studies, participants are presented with a narrative that provides misleading information about a previously experienced event. When participants are warned about the nature of the misleading information, they should become more skeptical about the origin of the postevent information. Therefore, the warning may allow them to be more resistant to the effects of the misleading information (Greene et al., 1982). Anastasi (1992) demonstrated that participants show higher recognition accuracy when they are given an explicit warning about the nature of the misleading information. In the current study,

an explicit warning was presented to participants before the recognition test. This warning may allow participants to be more cautious when determining whether an item was actually presented on the original list. Thus, the warning may lead participants to be resistant to the false memory effect by causing them to monitor the source of the test items more effectively. Therefore, the purpose of Experiment 2 was to determine whether participants are able to distinguish between actual memories and false memories if they are given an explicit warning.

Half the participants in Experiment 2 received an explicit warning explaining the nature of the false memory paradigm after the word list presentation but before the recognition test. The warning informed participants about the composition of the word lists, including information about the associative nature of the list. The remaining participants received no such warning. It was predicted that if participants could distinguish between the critical lures and the list items, they would be less likely to claim to remember the critical lures when given the warning. In addition to the warning, participants were asked to determine how confident they were that each word was presented previously. It was thought that these confidence judgments would provide a finer analysis of participants' memory than a simple yes/no recognition judgment. In other words, if participants were able to distinguish between the false memories and the actual memories, a confidence judgment would allow them to make a more precise memory decision than would a binary yes/no recognition judgment.

Roediger and McDermott (1995) have demonstrated that participants tend to be as confident that the critical lures were presented as list items. Therefore, we expected to replicate their findings that confidence ratings for the critical lures will be as high as the confidence ratings for the list items when no explicit warning was given. However, when participants were warned, the explicit warning was expected to lower the confidence ratings of the nonpresented critical lures in comparison to the list items. These predictions assume that participants can distinguish between the critical lures and the list items. If participants cannot make this distinction, then no difference would be expected in confidence ratings regardless of warning condition.

## METHOD

### Participants

Seventy-two students participated in the current study as partial credit for a library research requirement for an introductory psychology course at Francis Marion University. Participants were tested in groups of two to six. Half the

participants were given an explicit warning and the remaining participants were given no warning.

### Materials and procedure

The materials used in the current experiment were the same Nelson et al. (1994) word association lists used in Experiment 1. Unlike in Experiment 1, 24 word lists were used in Experiment 2 and were presented visually on a television screen in a 72-point Times New Roman font. Each item from 16 lists (224 items) was presented at a rate of 1.5 s per word (the remaining 8 lists were used to obtain distractors for the recognition test). Once the entire list was presented, participants were given the same word fragment completion task used in Experiment 1 and were told to complete as many of the word fragments as possible in the allotted 8 min.

After the word fragment completion task, an explicit warning was administered to half of the participants. Participants receiving the warning were told that the original word list contained 16 sublists. They were told that the sublists were devised so that each sublist was composed of associates to one particular central word. Therefore, the lists they saw were made up of items that were associated because they were all related to a central theme or central word. They were also told that the makeup of the study list often causes people to remember the central word even though that central word was not actually presented. Finally, before they were given the recognition test, participants were instructed to be very careful to indicate remembering only the items that were actually on the original list. The remaining participants received no such warning.

Participants were then given a 96-item memory test and were instructed to make confidence judgments about each of the words on the test using a 5-point scale. The confidence scale presented on the test consisted of 1 (*certain the item was presented*), 2 (*thought the item was presented*), 3 (*guessing*), 4 (*thought the item was not presented*), and 5 (*certain the item was not presented*). Participants were asked to use the entire scale in order to demonstrate the quality of their memory.

### RESULTS AND DISCUSSION

Experiment 2 allowed us to determine whether a false memory effect was found even when an explicit warning was given to participants. Table 1 shows the confidence ratings for each item type for both the warned and not-warned conditions. A 3 (list, critical lures, nonlist distractor)  $\times$

Table 1. Experiment 2: Confidence ratings for each item type and warning condition

Warning condition	List items	Critical lures	Nonlist items
Not warned	2.1 (1.7)	2.1 (1.7)	3.5 (2.9)
Warned	2.2 (1.7)	2.4 (2.0)	3.4 (2.8)

Note: Numbers in parentheses are mean confidence ratings after all ratings of 3 (*guessing*) were dropped.

2 (warned, not warned) ANOVA was conducted using the confidence rating data. Results revealed a main effect of item type,  $F(2, 140) = 219.80$ ,  $MSE = .18$ , but no main effect of warning,  $F(1, 70) = 1.31$ ,  $MSE = 0.50$ ,  $p = .26$ . However, there was a significant item type  $\times$  warning interaction,  $F(2, 140) = 6.57$ ,  $MSE = 0.18$ . Follow-up  $t$  tests revealed that the confidence rating for participants who were warned (2.2) and those who did not receive the warning (2.1) did not differ for the list items,  $t(36) = 0.82$ ,  $p = .42$ , or for the nonlist distractors (3.4 and 3.5 for the warned and not-warned conditions, respectively),  $t(36) = -0.97$ ,  $p = .34$ . However, participants were significantly more confident that the critical lures were presented on the original list when they received no warning (2.1) than when they received the explicit warning (2.4),  $t(36) = 2.57$ .

Additional follow-up  $t$  tests revealed that the critical lures and list items received the same confidence ratings for participants who did not receive the explicit warning (2.1 and 2.1, respectively),  $t(35) = .52$ ,  $p = .61$ . However, the confidence ratings for warned participants did reveal different ratings for the list items (2.1) and the critical lures (2.4),  $t(35) = -2.47$ ,  $p = .02$ . The critical lures were significantly different from the nonlist distractors under both the warned and the not-warned conditions,  $t(35) = 8.00$  and  $t(35) = 13.55$ , respectively.

One may argue that the confidence ratings used in the current experiment may be misleading because the ratings include a confidence rating of 3 (*guessing*). Therefore, analyses were also conducted by dropping all ratings of 3, which may bias the reported results. The data were rescored by using a 4-point confidence scale of 1 (*certain the item was presented*), 2 (*thought the item was presented*), 3 (*thought the item was not presented*), and 4 (*certain the item was not presented*). All differences were statistically identical, with very slight differences in the actual means. These means are presented in parentheses in Table 1.

It is important to note that although there was a statistically significant drop in confidence ratings for the critical lures once participants received a warning, the actual meaning of these numbers is important. Participants still believed that the critical lures were presented on the list, although they were less confident about these ratings. Therefore, the explicit warning did not eliminate the false memory effect but weakened it. In other words, participants seemed to be more resistant to false memories when they were given an explicit warning.

### EXPERIMENT 3

The final experiment used yet another phenomenological method to observe the nature of false memories. Based on the promising results of Experiment 2, a warning was used in the current experiment along

with another qualitative judgment (i.e., a remember/know judgment) that would allow us to better determine the phenomenological characteristics of the memory.

In addition to determining whether the items were "old" or "new" on the recognition test, participants were asked to make remember/know judgments concerning the items they recognized as being presented earlier in the experiment. Remember judgments are judgments in which participants have a specific episodic trace for the word presentation (Gardiner, 1988; Gardiner & Parkin, 1990; Payne et al., 1996; Rajaram, 1993; Tulving, 1985). In other words, participants may remember when the item was presented, what the item made them think of, words presented before or after the item, or how the item made them feel. In contrast, know judgments do not have this episodic component. When participants make a know judgment, they are also certain that the item was presented but they cannot remember the specific, episodic details of the presentation of the item (Gardiner, 1988; Gardiner & Parkin, 1990; Payne et al., 1996; Rajaram, 1993; Tulving, 1985). As stated earlier, it was expected that these judgments would allow a more precise estimate of the quality of participants' memory than a recognition judgment alone.

Gardiner and colleagues (Gardiner, 1988; Gardiner & Java, 1990; Gardiner & Parkin, 1990) performed several experiments demonstrating dissociations between remember and know judgments even though both of these judgments were made on items that were recognized as being presented earlier. These dissociations show that remember/know judgments do not simply provide the same information as recognition judgments. Gardiner (1988) demonstrated that the levels of processing effect and the generation effect occur only for remember judgments and are absent with know responses. Gardiner and Java (1990) found evidence of a word frequency effect for remember but not for know responses. In addition, Gardiner and Parkin (1990) showed that divided attention at study affected remember responses but not know responses. Levels of processing, generation, word frequency, and divided attention effects were seen only for recognition that was accompanied by conscious experience (i.e., remember judgments). It therefore seems clear that remember/know judgments provide additional information about the memory beyond that provided by recognition judgments alone.

Roediger and McDermott (1995) showed that participants typically claim to remember the critical lures more often than they claim to know that they were presented earlier (see Payne et al., 1996, for similar findings). This seems counterintuitive because the critical lures were never presented. In other words, how can one have a specific, episodic trace

for something that never actually occurred? If participants are unable to distinguish between the critical lures and the list items, then one would expect to see more remember responses for the critical lures regardless of the warning condition. On the other hand, if participants are able to make the distinction between words that were and were not actually presented, then one would expect to find more remember judgments for list items and more know judgments for the critical lures.

The final new aspect of Experiment 3 was to investigate a testing strategy that may have been used in previous false memory studies using a recognition test. One may argue that the false memory effect found in previous studies using a recognition test is simply a result of the testing strategy. The nonlist distractors on previous recognition tests were taken from other unrelated categories (i.e., distractors) and some were the most common associates of each of the presented lists (i.e., critical lures). It is possible that participants were simply remembering the general gist of the to-be-remembered list items and not the exact words. If this strategy was used, participants would be expected to indicate that they remembered the critical lures and perform very well at rejecting other nonlist items that were not associates of *any* list items. One way to determine whether this strategy is the cause of the false memory effect is by using other-category and same-category distractors on the final recognition test. Therefore, in Experiment 3, participants were presented with only half of the list items at study (i.e., a list of seven items) and were tested using list items, critical lures, same-category nonlist distractors, and other-category nonlist distractors. If the false memory effect is a result of the testing strategy to select any word related to the originally presented category, then participants will select the list items, critical lures, and same-category items at the same rate. However, if this is not the case, then one would expect participants to remember the list items but not same-category distractors. The other-category nonlist distractors should be rejected easily in either case, as has been found in previous studies.

In summary, Experiment 3 used the same explicit warning that was used in Experiment 2. This warning allowed participants to be more skeptical about the item's origins. Experiment 3 also used remember/know judgments in order to obtain a more sensitive memory quality judgment. As in Experiment 2, the warning may not have had an effect on participants' recognition judgment because they may have felt that the critical lures were familiar. However, their memory for the critical lures may not have been of the same quality as the list items, and the remember/know judgments may reveal this distinction. Finally, Experiment 3 was devised to explore a testing strategy that could account for the false memory effect found in previous studies using recognition

tests. The recognition test in the current study contains both same-category and different-category distractors so that participants cannot simply remember the gist of the word lists. However, if participants are solely remembering the gist of the word lists, then one would expect false memories with both critical lures and same-category distractors.

## METHOD

### Participants

Sixty students enrolled in an introductory psychology course at the State University of New York at Binghamton participated in the current study as partial credit for a library research requirement. Participants were tested in groups of two to six. As in Experiment 2, half of the participants were given an explicit warning and the remaining participants were not.

### Materials and procedure

The current study used the same materials as the previous experiments with one exception. Only seven items from each list were presented as the to-be-remembered items. These seven items were selected from the 16 lists such that each of the groups of seven items had approximately the same usage rate in the English language according to Thorndike and Lorge (1944). The frequency ratings for each list are presented in Appendix B. The remaining seven items served as a pool to obtain same-category distractors for the recognition test. These items were counterbalanced so that each group of seven items served equally often as list items and as the pool of same-category distractors. The order of the words within each list was random, as was the ordering of the lists.

As in Experiment 2, the 16 lists were presented one immediately following the other at a rate of 1.5 s per word. The items presented during the study session were blocked so that seven items from each list were presented, followed immediately by the seven items from the next list, and so on until all 112 words were presented. Once all lists were presented, participants were given the same word fragment completion filler task presented in Experiments 1 and 2. Following the filler task, half the participants were given the same explicit warning that was used in Experiment 2 before taking the memory test. The remaining participants received no warning and proceeded directly to the memory test.

Participants were then given a 144-item recognition test (48 list words, 48 same-category nonlist distractors, 32 other-category nonlist distractors, and 16 critical lures). The 48 list words were made up of 3 items randomly selected from one half of each of the 16 seven-item presented lists. The 48 same-category distractors were made up of 3 items randomly taken from the other, non-presented half of the 16 seven-item lists. The 32 novel distractors were made up of 3 randomly selected items from the 8 nonpresented lists (24 items) and the 8 critical lures for the nonpresented lists. The critical lures were the 16 critical lures corresponding to the 16 presented lists. For the *flower* list, an example of a list item was *blossom*, an example of a same-category nonlist item was

*daisy*, an example of a nonlist item was *beach*, and the critical lure was *flower*. Participants were instructed to write a "yes" next to the items they remembered being presented on the original study list and a "no" next to the items they felt were not presented earlier. In addition, participants were instructed to make remember/know judgments, which require a phenomenological judgment about the vividness of their memory for each word. For example, if they remembered other words presented before or after the word, if the word made them think of something else, or if they remembered how the word made them feel when it was presented, then they were instructed to write an *R* (*remember*) in the space provided. In other words, if they had a specific, episodic trace of the word being presented, then they were to write an *R*. However, if they were sure the item was presented but they were not able to specifically remember the particular episode or other details about the word, then they were instructed to write a *K* (*know*) in the space provided. Participants were instructed to make the remember/know judgment immediately after the yes/no judgment for each item.

## RESULTS AND DISCUSSION

### Recognition judgments

The results of Experiment 3 allowed us to determine whether false memories can be obtained even when same-category distractors are used and an explicit warning is given. A 4 (list item, same-category distractor, other-category distractor, critical lure)  $\times$  2 (warned, not warned) mixed-factor ANOVA was conducted. Means for each condition are shown in Table 2. The probability of calling an item "old" was no different for participants who were warned (.33) than for participants who were not warned (.38),  $F(1, 58) = 2.76$ ,  $MSE = 0.06$ ,  $p = .10$ . There was a main effect of item type,  $F(3, 174) = 119.36$ ,  $MSE = 0.02$ . More importantly, there was no significant item type  $\times$  warning interaction,  $F(3, 174) = 1.39$ ,  $MSE = 0.02$ ,  $p = .24$ . Planned comparisons showed that the probability of calling a same-category distractor "old" was lower when participants were given a warning (.25) than when no warning was given (.34),  $F(1, 58) = 4.28$ ,  $MSE = 0.03$ ,  $p = .04$ . In addition, participants called other-category distractors "old" less often when given a warning (.11) than when no warning was given (.17),  $F(1, 58) = 3.73$ ,  $MSE = 0.02$ ,  $p = .05$ .

Table 2. Experiment 3: Probability of calling an item "old" based on warning condition

Warning condition	List items	Critical lures	Same category		Other category		Average
No warning	.51	.51	.34	.17	.38		
Warning	.51	.45	.25	.11	.33		

However, there was no difference for the list items when a warning was given (.51) and when no warning was given (.51),  $F(1, 58) = 0.00$ ,  $MSE = 0.03$ ,  $p = .98$ , and there was also no difference in the proportion of critical lures recognized with the warning (.45) and without the warning (.51),  $F(1, 58) = 1.66$ ,  $MSE = 0.04$ ,  $p = .20$ . Therefore, the warning had no effect on the number of critical lures participants claimed to remember from the original list or the list items but did have an effect on the nonlist distractors.

Participants claimed to recognize the critical lures as list items nearly as often as they claimed to remember items that were actually presented. This provides further evidence of the robustness of the false memory phenomenon. Although there was no effect of the warning on the number of critical lures participants claimed to remember, it is possible that the quality of the memories was affected by the warning. In other words, it is possible that participants would make fewer remember judgments for the critical lures when they were given the explicit warning than when no warning was given. The warning may not affect the quantity of lures participants claimed to remember, but it may affect the quality of those memories.

#### Remember/Know judgments

In order to investigate the quality of participants' memories using the remember/know judgments, two separate 4 (list item, same-category distractor, other-category distractor, critical lure)  $\times$  2 (remember, know) within-subject ANOVAs were conducted: one for participants who were warned and one for participants who were not warned. The proportion of remember/know judgments that were remember judgments and the proportion of judgments that were know judgments were investigated using these data. The critical aspect of these data was the pattern of responses for each item type.

The first 4 (list, same category, other category, critical lure)  $\times$  2 (remember, know) ANOVA was conducted for participants who were not given any warning. The proportion of remember/know judgments that were remember responses and the proportion of judgments that were know responses are presented in the top half of Table 3. Results showed no main effect of item type,  $F(3, 87) = .66$ ,  $MSE = .00$ ,  $p = .57$ , or judgment,  $F(1, 29) = .13$ ,  $MSE = .35$ ,  $p = .72$ . More importantly, there was a significant item type  $\times$  judgment interaction,  $F(3, 87) = 18.97$ ,  $MSE = .07$ . When participants were not given a warning, the critical lures revealed a pattern similar to the list items (i.e., more remember than know responses) whereas the distractors showed more know than remember judgments. Simple effect tests showed that there were more remember judgments (.62) than know judgments (.38) for the list items,  $F(1, 29)$

Table 3. Experiment 3: Proportion of remember/know judgments for each item type and warning condition

Judgment	List items	Critical lures	Same category	Other category
		No warning		
Remember	.62	.59	.43	.30
Know	.38	.41	.57	.70
		Warning		
Remember	.55	.36	.36	.21
Know	.45	.64	.64	.79

$= 8.18$ ,  $MSE = .10$ . There were also more remember judgments (.59) than know judgments (.41) for the critical lures, although this difference was only marginally significant,  $F(1, 29) = 3.75$ ,  $MSE = .13$ ,  $p = .06$ . The opposite pattern was observed for the nonlist distractors. There was a significantly higher proportion of know judgments (.70) than remember judgments (.30) for the other-category distractors,  $F(1, 29) = 12.79$ ,  $MSE = .18$ , and there were more know judgments (.57) than remember judgments (.43) for the same-category distractors, although this difference was not significant,  $F(1, 29) = 1.71$ ,  $MSE = .15$ ,  $p = .20$ . Based on these data, the critical lures seem very similar to the list items and very different from the nonlist distractors. We now turn to the data for participants who were given the warning.

Another 4 (list item, same-category distractor, other-category distractor, critical lure)  $\times$  2 (remember, know) within-subject ANOVA was conducted for participants who received an explicit warning. These data are shown in the bottom half of Table 3. Results showed no main effect of item type,  $F(3, 87) = .73$ ,  $MSE = .00$ ,  $p > .53$ . However, there was a greater proportion of know judgments (.63) than remember judgments (.37),  $F(1, 29) = 15.10$ ,  $MSE = .27$ . More importantly, there was an item type  $\times$  judgment interaction,  $F(3, 87) = 19.29$ ,  $MSE = .06$ . Simple effect tests revealed a different pattern of remember/know judgments for the list items and the critical lures and a similar pattern for the critical lures and distractors. The list items showed more remember (.55) responses than know responses (.45), although this difference was not significant,  $F(1, 29) = 1.01$ ,  $MSE = .14$ ,  $p = .32$ . However, there were more know responses (.79) than remember responses (.21) for the other-category distractors,  $F(1, 29) = 92.56$ ,  $MSE = .06$ , as well as more know responses (.64) than remember responses (.36) for the same-category distractors,  $F(1, 29) = 9.80$ ,  $MSE = .12$ . Rather than obtaining more remember responses than know responses for the critical lures (as was found when no warning was given), we found that there were more know responses (.64) than remember responses (.36) for the critical lures,  $F(1, 29) =$

8.00,  $MSE = .14$ . This reversal demonstrates an effect of the warning that was not found with the recognition accuracy data.

The results of Experiment 3 showed that participants recognized the critical lures even when they were provided with an explicit warning specifying the robust nature of false memories. In addition, the current recognition results indicate that participants may be using a testing strategy to remember the list items. Recognition for the same-category distractors was significantly higher than for other-category distractors, but it was not as strong as for the list items or critical lures. This indicates that participants are using some sort of gist strategy. However, such a strategy does not fully account for the false memory effect because the critical lures are recognized at a much higher rate than the same-category distractors.

It is interesting to note that even when participants were warned about the nature of the lists, they still claimed to recognize the critical lures as items that were previously presented. However, the quality of the memory was affected by the warning instructions. There were different patterns of remember/know judgments for the critical lures depending on the warning condition. When participants were not given a warning, they produced more remember than know judgments for the critical lures, a pattern similar to that observed for the list items. However, when participants were given the explicit warning, they made more know judgments than remember judgments about the critical lures, a pattern similar to that observed for the other nonpresented distractors. Based on these results, it seems that the quality of the recollective experience is affected by the warning, even though this difference is masked by the potentially less sensitive recognition test.

### GENERAL DISCUSSION

The purpose of the current study was to determine whether participants could distinguish between memories of events that did actually occur and events that did not actually occur. Regardless of the warning condition, participants always believed that the critical lures were familiar and claimed to recognize the critical lures. If one were to simply measure the false memory effect with a recognition test or with other quality judgments, one would assume that participants could not distinguish between actual and false memories. Therefore, the current study used several methods that required participants to focus on the quality of their memory. Although familiarity seems to be the critical component with a yes/no recognition test, a deeper, qualitative judgment is required in order to make source monitoring judgments, confidence ratings, and remember/know judgments (Johnson, Kounios, & Reed-

er, 1994) especially when paired with an explicit warning. In addition, when participants are required to make judgments about the quality of the memory and given an explicit warning, it seems reasonable, based on the current findings, that they are able to distinguish between the phenomenological quality of the false and actual memories.

The implications of these findings are important within the context of false memory research as well as within the realm of our understanding of human memory in general. The false memory research illustrates the fact that human memory is not verbatim or exact (Loftus, 1997; Lynn & Payne, 1997; Mather et al., 1997; Payne et al., 1997; Roediger & McDermott, 1995), but rather is typically a distortion or gist representation (Loftus, 1997; Mather et al., 1997; Norman & Schacter, 1997; Payne et al., 1996).

Several related ideas have been proposed that may account for these memory distortions found in the false memory research. For example, Underwood (1965) reliably produced false memories and showed the importance of the associations between words. Underwood (1965) reasoned that when each word was initially presented, participants implicitly activated all words that were associates of the presented word. This was called an implicit associative response (IAR). Underwood concluded that words with very strong associates would be more likely to reveal a false memory effect. For instance, in false memory experiments participants may be presented with several words that are related to *sleep* but are never presented with the word *sleep*. When they are asked to remember the sleep-related list items, they remembered all of the activated words including the nonpresented word *sleep*. This simple explanation is similar to the idea behind the associative networks proposed by Collins and Quillian (1972). One common assumption when working with these associative networks is that when one item is activated, all other associated items are also activated. In addition, items that are stronger associates are activated more than weaker associates. It stands to reason that the critical lures used in the false memory studies using the Roediger and McDermott (1995) paradigm could be viewed as the strongest associates of the list items because the lists were devised based on those central, critical lures.

Another related theory that explains the false memory phenomenon is reality monitoring (Johnson, 1988; Johnson, Foley, & Leach, 1988; Johnson et al., 1994; Johnson & Raye, 1981). Reality monitoring theory, much like Underwood's IAR, also assumes that some sort of activation during encoding takes place. However, the false memories occur because participants are unable to determine whether the critical lure was actually presented or they merely thought of it during the encoding phase (Johnson, Raye, Foley, & Foley, 1981; Lane & Zaragoza, 1995;

Seamon et al., 1998). Thus, the problem becomes one of reality monitoring (Johnson, 1988; Johnson et al., 1988; Johnson, et al., 1994; Johnson & Raye, 1981). In other words, if participants are able to determine whether they actually saw an item or they merely thought it, then one would expect no false memories. However, in reference to the current findings, it appears that participants cannot perform reality monitoring effectively, and subsequently generate false memories.

Two conditions in the current study specifically instructed participants to determine where certain information originated: the source monitoring task and the explicit warning. Although participants may have used some source monitoring on the recognition tests, albeit ineffectively, the source monitoring instructions required participants to monitor their memories and try to determine the source of the item. One problem with the source monitoring task used in the current study is that it was not identical to a reality monitoring task. Participants were not given the opportunity to indicate that they remembered thinking of the critical lures at encoding. Ongoing research in our laboratory provides participants with an explicit option on the memory test that allows them to report that they do not remember the item being presented in the experiment, but do remember thinking of the critical lure while being presented with other items. This simple change may provide participants with the option they need in order to distinguish between actual and false memories. The explicit warning also instructs participants to monitor the source of their memories. Although the warning was not effective in eliminating false memories using recognition judgments, it did help participants differentiate between true list items and false memories when using the more sensitive memory quality judgments.

The false memory effect has already attracted much attention from memory researchers (Begley, 1996; Ceci et al., 1994; Deese, 1959; Gallo et al., 1997; Garry, Loftus, & Brown, 1994; Lindsay & Kelley, 1996; Lindsay & Read, 1994; Lynn & Payne, 1997; Mather et al., 1997; Norman & Schacter, 1997; Payne et al., 1996; Payne et al., 1997; Pezdek & Roe, 1994; Read, 1996; Rhodes & Anastasi, in press; Roediger & McDermott, 1995; Saywitz & Moan-Hardie, 1994; Schacter et al., 1996; Schooler, 1994; Seamon et al., 1998). The future will undoubtedly lead to many more investigations of this intriguing phenomenon. Although Deese (1959) has demonstrated a procedure that seems to produce the effect reliably, the current study provides valuable methods and findings that may be helpful to determine the causes and applications of this phenomenon. In any case, the present study provides insight into the workings of memory in general and may aid clinicians, police officers, and others who rely on an individual's memory report.

### Appendix A. Word lists used in the current study

All words used as study items are listed. *Italic words* are the three words from each list that were used as test items on the final recognition test. **Bold words** (the first words in each list) are the critical lures that were not presented on the study list, but did serve as distractor items on the recognition test. All items in each list are associates to the critical lure for that list.

Baby	Bomb	Drunk	Flower	Island	Money	Movie	Music
child	<i>explode</i>	alcohol	<i>rose</i>	water	cash	film	listen
cry	war	<i>beer</i>	petals	tropical	<i>spend</i>	theater	song
mother	nuclear	drive	smell	sand	green	popcorn	note
infant	atomic	drink	plant	sun	dollar	show	tune
<i>bottle</i>	<i>threat</i>	sick	daisy	<i>beach</i>	wealth	picture	stereo
crib	death	<i>sober</i>	pot	sea	rich	<i>television</i>	radio
doll	shelter	liquor	garden	Hawaii	buy	screen	dance
little	destruction	bar	bloom	<i>paradise</i>	check	star	band
boy	Hiroshima	<i>intoxicate</i>	seed	ocean	bank	cinema	guitar
<i>girl</i>	<i>shell</i>	bum	blossom	palm	<i>dough</i>	<i>flak</i>	<i>mobility</i>
diaper	destroy	booze	bed	land	coin	reel	piano
rattle	blow	wine	vase	vacation	<i>currency</i>	director	<i>record</i>
carriage	fire	plastered	tulip	<i>deserted</i>	luxury	video	sound
<i>kid</i>	boom	inebriated	<i>stem</i>	resort	bill	<i>actor</i>	sing

Planets	Rabbit	River	Sew	Wind	Witness	Wood	Writer
Earth	<i>bunny</i>	<i>lake</i>	<i>needle</i>	<i>blow</i>	<i>crime</i>	tree	author
<i>solar</i>	fur	stream	thread	breeze	court	burn	book
<i>astronomy</i>	white	<i>water</i>	clothes	<i>rain</i>	see	<i>forest</i>	typewriter
space	<i>animal</i>	canoe	button	air	jury	chop	<i>poet</i>
sun	<i>ears</i>	boat	knit	cool	trial	log	novel
Mars	jump	bend	stitch	storm	<i>saw</i>	fireplace	reader
universe	hare	bank	<i>tear</i>	cold	eye	oak	<i>pen</i>
moon	hunt	Mississippi	<i>cloth</i>	kite	<i>testify</i>	splinter	composer
galaxy	<i>Easter</i>	rapids	mend	mill	murder	bark	editor
<i>Saturn</i>	stew	<i>ocean</i>	yarn	<i>surf</i>	person	natural	script
Pluto	carrot	brook	seam	hurricane	stand	cut	journalist
Mercury	soft	swim	hem	Chicago	defendant	<i>build</i>	reporter
orbit	cottontail	<i>cool</i>	fabric	tornado	testimony	<i>board</i>	publisher
Jupiter	hop	fish	pin	<i>snow</i>	judge	saw	story

Soup	Stomach	Office	Lunch	Mouth	Diamond	Airplane	Chicken
<i>tomato</i>	<i>ache</i>	<i>building</i>	food	<i>teeth</i>	<i>expensive</i>	<i>helicopter</i>	eggs
chicken	<i>food</i>	principal	breakfast	wash	emerald	flight	rooster
hot	bile	work	dinner	lips	ring	airport	hen
Campbell's	hungry	computer	<i>school</i>	open	necklace	airline	<i>barbecue</i>
vegetables	intestine	executive	break	gums	rock	jet	hatch
<i>baoul</i>	tummy	business	sandwich	talk	engagement	crow	McNuggets
<i>sharp</i>	belly	<i>secretary</i>	supper	<i>kiss</i>	<i>sportlike</i>	<i>wings</i>	

Continued on next page

Appendix A. *Continued*

Soup	Stomach	Office	Lunch	Mouth	Diamond	Airplane	Chicken
broth	abdomen	assistant	meal	spit	ruby	departure	chuck
cracker	<i>organ</i>	desk	eat	tongue	earring	ride	bird
spoon	esophagus	doctor	noon	sativa	gem	fly	fried
salad	acid	room	<i>afternoon</i>	nose	gold	trip	<i>farm</i>
<i>stew</i>	ulcer	supplies	snack	speak	pearl	stewardess	wing
can	fat	chairs	hour	breath	hard	pilot	duck
chowder	upset	<i>telephone</i>	bag	<i>face</i>	<i>jewel</i>	<i>aircraft</i>	feathers

## Appendix B. Frequency ratings for each list in Experiment 3

The average frequency rating according to Thorndike and Lorge (1944) for each list used in Experiment 3 is listed. The list name as well as the frequency rating is listed for both halves of the list. The average of each list is presented so that a comparison of list halves can be compared.

	List A1	List A2
Rabbit	49	50
Planets	39	39
Flower	52	51
Drunk	43	43
Baby	61	61
Office	76	71
Soup	30	33
Stomach	30	29
Average	47	47

  

	List B1	List B2
Writer	39	39
Eyewitness	68	63
Lunch	79	75
Mouth	68	72
Bomb	56	58
Music	66	68
Wood	65	64
Diamond	46	43
Average	61	60

  

	List C1	List C2
Windy	64	65
Airplane	43	42
Chicken	45	51
River	68	67

*Continued on next page*

Appendix B. *Continued*

	List C1	List C2
Movie	40	39
Island	55	52
Sew	39	40
Money	73	71
Average	53	53

## Notes

We thank David G. Payne, Michael J. Wenger, Marcia K. Johnson, J. Don Read, Donaldson E. Dulany, Cheryl Sterling, John Hester, and an anonymous reviewer for their comments on early versions of this manuscript. We also thank Julie Golden, Meghan Murphy, Stacy Barnett, Molly Zebraski, and Dan Sloat for their assistance with the experiments reported here. Correspondence concerning this article should be addressed to Jeffrey S. Anastasi, Department of Psychology, Francis Marion University, Florence, SC 29505-0547 (e-mail: janastasi@fmarion.edu). Received for publication March 12, 1997; revision received May 19, 1998.

## References

- Anastasi, J. S. (1992). *Explicit warnings decrease the misleading post-event suggestibility effect*. Unpublished manuscript.
- Anastasi, J. S., Burns, M. C., & Golden, J. (1996, March). *False memories: Are they indistinguishable from actual memories?* Poster presented to the 42nd Annual Meeting of the Southeastern Psychological Association, Norfolk, VA.
- Begley, S. (1996, July 15). You must remember this: but do you really? *Newsweek*, 64.
- Belli, R. F., Windschitl, P. D., McCarthy, T. T., & Winfrey, S. E. (1992). Detecting memory impairment with a modified test procedure: Manipulating retention interval with centrally presented event items. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 356-367.
- Ceci, S. J., Huffman, M. L., & Smith, E. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition*, 3, 388-407.
- Chandler, C. G. (1989). Specific retroactive interference in modified recognition tests: Evidence for an unknown cause of interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 256-265.
- Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and post-event information. *Memory & Cognition*, 11, 467-475.
- Collins, A. M., & Quillian, M. R. (1972). How to make a language user. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 309-351). New York: Academic Press.

*Continued on next page*

- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17-22.
- Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, *4*, 271-276.
- Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309-313.
- Gardiner, J. M., & Java, R. I. (1990). Recognizing and remembering. In A. Collins, S. Gathercole, & P. Morris (Eds.), *Theories of memory* (pp. 168-188). Hillsdale, NJ: Erlbaum.
- Gardiner, J. M. & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, *18*, 579-583.
- Garry, M., Loftus, E. F., & Brown, S. W. (1994). Memory: A river runs through it. *Consciousness and Cognition*, *3*, 438-451.
- Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior*, *21*, 207-219.
- Johnson, M. K. (1988). Discriminating the origin of information. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs* (pp. 34-65). New York: Wiley.
- Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory & Cognition*, *16*, 337-342.
- Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1409-1419.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67-85.
- Johnson, M. K., Raye, C. L., Foley, J. J., & Foley, M. A. (1981). Cognitive operations and decisions bias in reality monitoring. *American Journal of Psychology*, *94*, 37-64.
- Lane, S. M., & Zaragoza, M. S. (1995). The recollective experience of cross-modality confusion errors. *Memory & Cognition*, *23*, 607-610.
- Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, *17*, 349-358.
- Lindsay, D. S., & Kelley, C. M. (1996). Creating illusions of familiarity in a cued recall remember/know paradigm. *Journal of Memory and Language*, *35*, 197-211.
- Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology*, *8*, 281-338.
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560-572.
- Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
- Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist*, *48*, 518-537.
- Loftus, E. F. (1997). Memory for a past that never was. *Current Directions in Psychological Science*, *6*, 60-65.

## MEMORY ILLUSIONS

- Loftus, E. F., Donders, K., Hoffman, H., & Schooler, J. W. (1989). Creating new memories that are quickly accessed and confidently held. *Memory & Cognition*, *17*, 607-616.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585-589.
- Lynn, S. J., & Payne, D. G. (1997). Memory as the theater of the past: The psychology of false memories. *Current Directions in Psychological Science*, *6*, 55.
- Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/Know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, *25*, 826-837.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1994). *The University of South Florida word association, rhyme, and word fragment norms*. Unpublished manuscript.
- Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, *25*, 838-848.
- 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.
- Payne, D. G., Neuschatz, J. S., Lampinen, J. M., & Lynn, S. J. (1997). Compelling memory illusions: The qualitative characteristics of false memories. *Current Directions in Psychological Science*, *6*, 56-60.
- Payne, D. G., Toggia, M. P., & Anastasi, J. S. (1994). Recognition performance level and the magnitude of the misinformation effect in eyewitness memory. *Psychonomic Bulletin & Review*, *1*, 376-382.
- Pezdek, K., & Roe, C. (1994). Memory for childhood events: How suggestible is it? *Consciousness and Cognition*, *3*, 374-387.
- Piaget, J. (1951). *Plays, dreams, and imitation in childhood*. New York: Norton.
- Pynoos, R. S., & Nader, K. (1989). Children's memory and proximity to violence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*, 236-241.
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & 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.
- Rhodes, M. G., & Anastasi, J. S. (in press). The effects of a levels of processing manipulation on false recall. *Psychonomic Bulletin & Review*.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803-814.
- Sayewitz, K. J., & Moan-Hardie, S. (1994). Reducing the potential for distortion of childhood memories. *Consciousness and Cognition*, *3*, 408-425.
- Schacter, D. L. (1996). *Searching for memory* (pp. 248-279). New York: Basic Books.

- Schacter, D. L., Reiman, E., Curran, T., Yun, L. S., Bandy, D., McDermott, K. B., & Roediger, H. L., III. (1996). Neuroanatomical correlates of veridical and illusory recognition memory revealed by PET. *Neuron, 17*, 267-274.
- Schooler, J. W. (1994). Seeking the core: The issues and evidence surrounding recovered accounts of sexual trauma. *Consciousness and Cognition, 3*, 452-469.
- 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.
- Smith, V. L., Kassin, S. M., & Ellsworth, P. C. (1989). Eyewitness accuracy and confidence: Within- versus between-subjects correlations. *Journal of Applied Psychology, 74*, 356-359.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University Press.
- Toglia, M. P., Ross, D. F., Ceci, S. J., & Hembrooke, H. (1992). The suggestibility of children's memory: A social-psychological and cognitive interpretation. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 217-241). New York: Springer-Verlag.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1-12.
- Tulving, E. L., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 336-342.
- Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122-129.
- Zaragoza, M. S., & Koshmider, J. W., III. (1989). Mised participants may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 246-255.