The effects of a levels-of-processing manipulation on false recall

MATTHEW G. RHODES and JEFFREY S. ANASTASI
Francis Marion University, Florence, South Carolina

The present study attempted to determine the effect of a levels-of-processing manipulation on the incidence of false recall. In Experiment 1, participants engaged in either a vowel counting task or a concrete/abstract rating task; in Experiment 2, participants engaged in either a vowel counting task or a category sorting task. Results of both experiments demonstrated that participants who engaged in a deeper level of processing (i.e., concrete/abstract ratings or category sorting) recalled significantly more list items and critical lures. The present findings thus lend support to theories that attribute false memories to activation-based factors.

The subject of false memories—memories for events that never occurred—has received a great deal of attention in the past few years. Roediger and McDermott (1995) discussed a long-ignored study conducted by Deese (1959) that provides a method that consistently produces false memories. Deese presented participants with a list of words that were associates of a single, nonpresented item. For example, participants were presented with a list of words that are the twelve most common associates to the word *sleep*, such as *bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore,* and *nap,* but the participants were never presented with the critical lure *sleep.* Results demonstrated that participants consistently recalled the critical lure from each list at a level comparable to the list items. Deese's study therefore provides a method by which false memories can be reliably produced. Numerous studies have since replicated Deese's findings and have found the paradigm to produce a robust false-memory effect (Anastasi, Rhodes, & Burns, 2000; Gallo, Roberts, & Seamon, 1997; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996; Payne, Neuschatz, Lampinen, & Lynn, 1997; Roediger & McDermott, 1995; Seamon, Luo, & Gallo, 1998). However, much of the current literature has focused on the boundary conditions sufficient for producing false memories but has ignored the underlying processes and theoretical factors involved in the creation of false memories.

Underwood's (1965) implicit associative response (IAR) theory is the most often cited explanation for the false-memory phenomenon. Underwood proposed that when a word is presented during encoding, a participant may implicitly activate an associate of that word. For example,

when the word *day* is presented, a participant may think of the associate *night.* At retrieval, participants may recognize or recall an associate because it had been activated either implicitly or explicitly at encoding regardless of whether the associate was ever actually presented. Other more recent theories have further elaborated Underwood's semantic activation-based approach (e.g., Ayers & Reder, 1998).

Ayers and Reder's (1998) source of activation confusion (SAC) model may provide the most salient explanation of the factors affecting semantic activation. In essence, SAC is based on the traditional associationist memory models (Anderson, 1976, 1993; Collins & Loftus, 1975; Collins & Quillian, 1972; Sternberg, 1996). The major premise of SAC is that both the degree to which a concept is activated and the number of times it is activated will increase the strength of the associative network and result in better memory for that item. For example, viewing such words as *bed, rest, awake,* and *pillow* will lead to a strengthening of the associations between each of these items as well as other related items, such as the critical lure *sleep.* In addition, deeper processing or multiple presentations of list items will result in more extensive activation and, subsequently, a greater rate of recall for the presented items. The SAC model would also posit that the increased activation that aided memory for list items would likewise lead to an increase in recall for the critical lures that were never presented.

In contrast, the levels-of-processing approach (Craik & Lockhart, 1972) would make differing predictions concerning the recall of nonpresented associates. The concept of levels of processing was first introduced by Craik and Lockhart, who contend that the retentive value of an item is a function of the level of processing by which the item is encoded. In other words, "deeper" cognitive processing of an item at encoding (e.g., defining the item or producing associates of the item) will lead to very accurate recollection. Conversely, "shallow" cognitive processing (e.g., determining how many letters are in a par-

ticular word or determining whether the word has any vowels) will lead to poorer memory performance. Many studies have since supported this idea that deeper processing of presented items leads to greater retention (Craik & Tulving, 1975; Fisher & Craik, 1977, 1980; Flannagan & Blick, 1989; Toth, 1996; Vochatzer & Blick, 1989). Hence, the concept has come to be regarded as one of the most well-established information processing theories of the past 20 years (Lockhart & Craik, 1990). In spite of this wide acceptance, very little research has attempted to relate levels of processing to the false-memory literature (see Read, 1996, and Tussing & Greene, 1997, for exceptions), while it seems as though a manipulation that reliably increases memory performance would be crucial when investigating the relationship between veridical memories and memory illusions. Essentially, deeper processing of list items should result in more accurate overall recall and, consequently, better memory of the presented items and fewer nonlist intrusions.

Read (1996, Experiment 2) attempted to address the relationship between levels of processing and false memories by instructing participants to use one of three encoding strategies that differed in terms of the level of processing required. The levels-of-processing conditions were (from deepest to most shallow) elaborative rehearsal, maintenance rehearsal, and serial rehearsal. Participants in the elaborative rehearsal condition were told to "think about and rehearse the words in ways that would allow them to later answer questions about word meanings" (Read, 1996, p. 107). Participants employing maintenance rehearsal were instructed to always keep the last word in mind. Participants in the serial rehearsal condition were told to remember the words in the exact order that they were presented. However, one of the major shortcomings of the study is that these different levels-of-processing conditions are similar and defined rather ambiguously. For example, it is difficult to say precisely what level of processing participants in the elaborative rehearsal condition utilized. Additionally, it is difficult to argue for meaningful differences between participants who utilized maintenance and serial rehearsal. Read's results support this idea that the differences between the levels of processing used at encoding are ambiguous. Participants who engaged in the deepest level of processing (i.e., elaborative rehearsal) recalled critical lures at the same rate as those who utilized maintenance rehearsal at encoding. However, the rates of false recall for participants in the elaborative and maintenance rehearsal conditions were significantly higher than for those who encoded on the most shallow level (i.e., serial rehearsal). Nevertheless, on the basis of the encoding instructions, it is unclear why these differences were found since it is difficult to discern the qualitative differences that distinguish the three encoding conditions.

Tussing and Greene (1997) also investigated the false-memory effect using a levels-of-processing paradigm. In a manner comparable to that in the Read (1996) study, Tussing and Greene utilized three levels of processing. Participants were required to determine whether the first letter of each presented word was a vowel (their most shallow level of processing), count the number of letters in each word, or make a 7-point pleasantness rating as each word was presented (their deepest level of processing). Results from their recognition data indicate that participants in the letter counting condition remembered the lowest proportion of critical lures (.73), whereas participants in the vowel judgment condition (.81) and pleasantness rating condition (.82) remembered critical lures at virtually identical rates. One major problem with interpreting these results is the fact that there were no differences in recognition accuracy between the levels-of-processing conditions for the list items. Therefore, as noted by Tussing and Greene, it is difficult to interpret the critical lure data since they were unable to demonstrate a levels-of-processing manipulation with the list items.

In an earlier study conducted in our lab (Anastasi, Rhodes, Carter, & Gaddy, 1998), participants' memory for list items and critical lures was examined by utilizing two different levels of processing. Participants using a shallow level of processing were instructed to count the number of vowels in each list item as it was presented, and participants using a deep level of processing were instructed to name an associate of each list item as it was presented. Results demonstrated that the levels-of-processing manipulation had a significant impact on the proportion of list items recalled (.13 and .31 for the shallow and deep conditions, respectively) as well as the proportion of critical lures recalled (.25 and .75 for the shallow and deep conditions, respectively). However, an alternative explanation for these findings may be derived from the deep-level-of-processing task that was used. In several instances when participants were asked to write down the first word that came to mind, they wrote down the critical lure. Thus, it may be argued that participants were merely recalling the associate that they had written down during encoding (i.e., the critical lure). Although this study provides evidence that a deeper level of activation leads to increased recall of critical lures, further investigation of the matter utilizing alternative deeper levels of processing is necessitated.

EXPERIMENT 1

In Experiment 1, we investigated the effect of levels of processing on false memories by using well-defined and qualitatively different encoding manipulations. Specifically, participants were required to either count the number of vowels in each word as it was presented (e.g., the shallow-level-of-processing condition) or make a concrete/abstract rating for each item (e.g., the deep-level-of-processing condition).

In accordance with the levels-of-processing theory, participants who encode the word list using a deeper level of processing should recall a greater proportion of list items than participants who encode the list items using a shallow level of processing. Moreover, this deeper level of processing should result in more accurate memory and, consequently, fewer memory illusions. Conversely, an

activation-based approach would posit that the more deeply an item is processed, the greater the activation between the list items and their associates. For example, encoding such items as *bed, rest, awake, pillow*, and *dream* using a deep level of processing will lead to a strengthening of the associations between these items as well as nonpresented, related items, such as the critical item *sleep*.

In sum, a levels-of-processing approach would predict better memory for list items and few memory illusions for the deep-processing condition, whereas an activation-based approach would predict better memory for the list items and a greater incidence of false memories. It thus stands to reason that if participants recall fewer critical lures in the deep-level-of-processing condition, then the activation-based models do not provide an adequate theoretical explanation for the false-memory phenomenon. On the other hand, greater recall for critical lures in the deep-processing condition would provide support for the activation-based models.

Method

Participants. Forty Francis Marion University introductory psychology students participated as part of a class requirement for research participation. The participants were run in groups ranging from 2 to 5 individuals.

Design. A 2 (level of processing: shallow, deep) × 2 (item type: list, critical lure) mixed-factor design was used, with level of processing manipulated between subjects and item type manipulated within subjects. The dependent variable was the proportion of items recalled.

Materials and Procedure. The four 15-item sublists used in the present experiment were the *needle, sleep, rough*, and *soft* lists taken from Roediger and McDermott (1995). The sublists were ordered in such a manner that the strongest associates occurred earlier in the list, followed by the weaker associates. In addition, 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.

Each item was presented at a 3-sec rate using an audio tape player. The participants in the shallow-level-of-processing condition were instructed to count and write down the number of vowels in each word as it was presented. For example, the participants were told that if they were presented with the word *dog*, they would write down a "1," because there is one vowel in that word. The participants in the deep-level-of-processing condition were instructed to visualize each word as it was presented and then make a concrete/abstract rating. For example, the participants were told that *tree* is a very concrete item since one can touch it, whereas *justice* is very abstract since it is impossible to touch. These participants were given a scale on the top of a rating sheet that had a range of choices from 1 (*very concrete*) to 5 (*very abstract*). Following the list presentation, the participants engaged in a 5-min unrelated filler task. The purpose of this task was to eliminate any rehearsal of list items in addition to any recency effects of memory. Finally, the participants were administered a free-recall test and instructed to write down as many of the list items as they could remember in any order. They were given 5 min for the recall test.

Results and Discussion

A 2 (level of processing: shallow, deep) × 2 (item type: list, critical lure) analysis of variance (ANOVA) was conducted on the proportion of items recalled. Table 1 presents the mean proportions of list items and critical lures recalled for both the shallow- and deep-processing con-

Table 1
Proportions of List Items and Critical Lures Recalled in Experiment 1

Encoding Condition	Item Type	
	List	Critical
Shallow	.18	.23
Deep	.29	.47

ditions. Results revealed a main effect of both level of processing [$F(1,38) = 8.81$, $MS_e = 0.07$, $p < .05$ (an alpha level of .05 was used for all statistical tests unless otherwise noted)] and item type [$F(1,38) = 7.03$, $MS_e = 0.04$]. However, there was no level of processing × item type interaction [$F(1,38) = 2.00$, $MS_e = 0.04$]. Follow-up t tests were conducted to further examine these data. In line with the levels-of-processing theory, the participants in the deep-level-of-processing condition (.29) recalled a significantly greater proportion of list items than did the participants in the shallow-level-of-processing condition (.18) [$t(38) = -2.875$]. More importantly, participants in the deep-level-of-processing condition recalled a significantly greater proportion of critical lures than did the participants in the shallow-level-of-processing condition (.47 and .23, respectively) [$t(38) = -2.459$]. Hence, the present findings lend support to the idea that a deeper level of processing results in greater activation of list items and their associates rather than an overall improvement in memory performance as predicted by the levels-of-processing approach.

A follow-up t test was also conducted to determine whether the participants in the deep-level-of-processing condition demonstrated higher guessing rates than the participants in the shallow-level-of-processing condition using the number of noncritical intrusions recalled. Specifically, a greater number of noncritical intrusions in the deep-level-of-processing condition would indicate that the participants were employing a guessing strategy, the result of which would be greater recall of list items and critical lures as well as the noncritical instrusions. However, results indicated that there was no difference in the number of noncritical intrusions recalled between the deep-level-of-processing condition (3.95) and the shallow-level-of-processing condition (2.95) [$t(38) = -0.70$].

The results of Experiment 1 demonstrated that the deeper level of processing utilized at encoding led to greater activation of the list items and their associates, including the critical lures. This is illustrated by the fact that the levels-of-processing manipulation had a significant effect on the proportion of list items recalled as well as critical lures. Hence, a deeper level of processing led to more accurate memory for list items but, as activation-based theories would predict, a greater number of memory illusions.

EXPERIMENT 2

Experiment 2 was an attempt to replicate the results of Experiment 1 using an alternative deep-level-of-processing

condition. While the shallow-level task remained identical to that used in Experiment 1, the deep-level-of-processing task utilized a very different type of processing. Specifically, participants in the deep-level condition were instructed to perform a category sorting task as opposed to making the concrete/abstract judgments required in Experiment 1. Although these tasks involved very different types of processing, they can both be characterized as utilizing a deeper level of processing than the vowel counting task. Due to the nature of the category sorting task, some slight methodological modifications were necessary. Since participants in the deep-level-of-processing condition were required to categorize each presented item into one of four categories, the list items were presented in a mixed format rather than a blocked format. Past research has shown that a mixed format may decrease the level of false memories (Mather, Henkel, & Johnson, 1997; McDermott, 1996). However, what was important in Experiment 2 was the pattern of results and not a comparison across experiments. Despite these modifications, Experiment 2 was expected to replicate the pattern of results of Experiment 1.

Method

Participants. Forty Francis Marion University introductory psychology students participated as part of a class requirement for research participation. The participants were run in groups ranging from 2 to 4 individuals.

Design. A 2 (level of processing: shallow, deep) × 2 (item type: list, critical lure) mixed-factor design was used, with level of processing manipulated between subjects and item type manipulated within subjects. The dependent variable was the proportion of items recalled.

Materials and Procedure. The present experiment used the same word lists used in Experiment 1 with some slight methodological alterations. The participants were given 60 5 × 7 in. index cards, each of which had a 24-point font word typed in the center of the card. The participants in the shallow-level-of-processing condition were instructed to view the word on the index card, count the number of vowels, and write down that number on a sheet of paper provided. Participants in the deep-level-of-processing condition were given a sheet of paper that was divided into four boxes and were instructed to categorize the words on the index cards by writing them down in any one of the four boxes. The participants were told that they could categorize the words on the index cards in any manner that they felt was appropriate. The order of the index cards was random. All participants were told that once they had made their judgment, they could not view that index card again. The procedure for the filler task and recall test was identical to that of Experiment 1.

Results and Discussion

A 2 (level of processing: shallow, deep) × 2 (item type: list, critical lure) ANOVA was conducted on the proportion of items recalled. Table 2 presents the mean proportions of list items and critical lures recalled for both the shallow- and deep-processing conditions. Results revealed a main effect of level of processing [$F(1,38) = 61.88, MS_e = 0.03$], but no effect of item type [$F(1,38) = 0.29, MS_e = 0.02$]. In addition, there was no level of processing × item type interaction [$F(1,38) = 0.29, MS_e = 0.29$]. These data were further examined using follow-up t tests. Replicating Experiment 1, the participants in

Table 2
Proportions of List Items and Critical Lures Recalled in Experiment 2

Encoding Condition	Item Type	
	List	Critical
Shallow	.09	.09
Deep	.38	.41

the deep-level-of-processing condition a recalled significantly greater proportion of list items (.38) than did the participants in the shallow-level-of-processing condition (.09) [$t(38) = -9.55$]. More importantly, the participants in the deep-level-of-processing condition recalled a significantly greater proportion of critical lures than did the participants in the shallow-level-of-processing condition (.41 and .09, respectively) [$t(38) = -5.06$].

Finally, there was a significant difference in the number of noncritical intrusions recalled between the deep-level-of-processing condition (1.05) and the shallow-level-of-processing condition (2.80) [$t(38) = 2.37$]. Despite the fact that the participants in the shallow-level-of-processing condition exhibited more noncritical intrusions, they recalled far fewer list items and critical lures than the participants in the deep-level-of-processing condition.

GENERAL DISCUSSION

Consistent with the levels-of-processing framework and activation-based models, results from both experiments indicated that the participants in the deep-level-of-processing conditions (i.e., concrete/abstract ratings, category sorting) were significantly more likely to recall the list items than the participants in the shallow-level-of-processing conditions (i.e., counting vowels). More importantly, the participants in the deep-level-of-processing conditions recalled a significantly greater proportion of critical lures than those in the shallow-level-of-processing conditions. Thus, the levels-of-processing manipulation, while improving general memory performance, led to a greater number of memory illusions. This counterintuitive result cannot be explained solely by the levels-of-processing framework since fewer memory illusions in the deep-level-of-processing condition would be predicted. However, these findings can be easily explained by activation-based models. Encoding the list items using a deeper level of processing results in greater activation of the connections between presented items in the memory network, leading to greater recall of the list items and other related items, such as the critical lures. Therefore, the present findings lend greater veracity to theories explaining the false-memory phenomenon as the product of activation-based factors, an idea that parallels both Underwood's (1965) IAR theory and Ayers and Reder's (1998) SAC model. Arndt and Hirshman (1998) provide a more precise activation-based explanation. They contend that list items are recollected on the basis of the activation of a single memory trace, whereas critical lures

are remembered on the basis of the summation of several weaker associated traces. According to this reasoning, one may argue that veridical memories and memory illusions would then increase concurrently.

In reference to such activation-based theories, the most appropriate way to view false memories may be as a function of the level of activation that is elicited by the particular encoding task. This is not to say that retrieval is a negligible factor, quite the contrary. Underwood's (1965) IAR theory posits that the implicit activation of associates can occur both at encoding and during the act of retrieval. However, it is apparent that the level or strength of activation that is the result of the particular encoding task has a significant effect on the rate of false recall. Invariably, this is achieved through an elaboration and strengthening of the associative network and its related concepts. The present study clearly supports this idea. The shallow-level-of-processing condition led to very weak activation and, consequently, a rather meager associative network. This is demonstrated by the recall of fewer list items and other related items. On the other hand, a deep level of processing led to stronger connections in the associative network, which is illustrated by the fact that the participants in the deep-level-of-processing condition recalled list items and critical lures at rates significantly greater than the shallow-level-of-processing condition. This greater activation is not strictly a result of the level of processing but rather the level of activation that is brought about by the processing tasks employed at encoding.

Overall, the present study serves to add to a body of research that has demonstrated that false memories can be reliably produced and provides support for the idea that memory illusions are the result of semantic activation. Future studies may utilize the level-of-activation framework discussed here and use other encoding tasks that further manipulate the levels of activation to determine their effect on the incidence of memory illusions.

REFERENCES

ANASTASI, J. S., RHODES, M. G., & BURNS, M. C. (2000). Distinguishing between memory illusions and actual memories utilizing phenomenological measurements and explicit warnings. American Journal of Psychology, 113, 1-26.

ANASTASI, J. S., RHODES, M. G., CARTER, J. S., & GADDY, J. R. (1998, March). False memories: Encoding or retrieval factors? Poster presented at the 44th Annual Meeting of the Southeastern Psychological Association, Mobile, AL.

ANDERSON, J. R. (1976). Language, memory, and thought. Hillsdale, NJ: Erlbaum.

ANDERSON, J. R. (1993). Rules of the mind. Hillsdale, NJ: Erlbaum.

ARNDT, J., & HIRSHMAN, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. Journal of Memory & Language, 39, 371-391.

AYERS, M. S., & REDER, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. Psychonomic Bulletin & Review, 5, 1-21.

COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82, 407-428.

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.

CRAIK, F. I. M., & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning & Verbal Behavior, 11, 671-684.

CRAIK, F. I. M., & TULVING, E. (1975). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

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

FISHER, R. P., & CRAIK, F. I. M. (1977). Interaction between encoding and retrieval operations in cued recall. Journal of Experimental Psychology: Human Learning & Memory, 3, 701-711.

FISHER, R. P., & CRAIK, F. I. M. (1980). The effects of elaboration on recognition memory. Memory & Cognition, 8, 400-404.

FLANNAGAN, D. A., & BLICK, K. A. (1989). Levels of processing and the retention of word meanings. Perceptual & Motor Skills, 68, 1123-1128.

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.

LOCKHART, R. S., & CRAIK, F. I. (1990). Levels of processing: A retrospective commentary on a framework for memory research. Canadian Journal of Psychology, 44, 87-112.

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.

McDERMOTT, K. B. (1996). The persistence of false memories in list recall. Journal of Memory & Language, 35, 212-280.

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 & 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.

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.

ROEDIGER, H. L., III, & McDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, & Cognition, 21, 803-814.

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.

STERNBERG, R. J. (1996). Cognitive psychology. Fort Worth, TX: Harcourt Brace.

TOTH, J. P. (1996). Conceptual automaticity in recognition memory: Levels of processing effects on familiarity. Canadian Journal of Experimental Psychology, 50, 123-138.

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

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

VOCHATZER, K. G., & BLICK, K. A. (1989) Levels of processing and the retention of paired-associates. Perceptual & Motor Skills, 69, 349-350.