The importance of material–processing interactions in inducing false memories

JASON C. K. CHAN, KATHLEEN B. McDERMOTT, JASON M. WATSON, and DAVID A. GALLO
Washington University, St. Louis, Missouri

Deep encoding, relative to shallow encoding, has been shown to increase the probability of false memories in the Deese/Roediger–McDermott (DRM) paradigm (Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999). In two experiments, we showed important limitations on the generalizability of this phenomenon; these limitations are clearly predicted by existing theories regarding the mechanisms underlying such false memories (e.g., Roediger, Watson, McDermott, & Gallo, 2001). Specifically, asking subjects to attend to phonological relations among lists of phonologically associated words (e.g., *weep*, *steep*, etc.) increased the likelihood of false recall (Experiment 1) and false recognition (Experiment 2) of a related, nonpresented associate (e.g., *sleep*), relative to a condition in which subjects attended to meaningful relations among the words. These findings occurred along with a replication of prior findings (i.e., a semantic encoding task, relative to a phonological encoding task, enhanced the likelihood of false memory arising from a list of semantically associated words), and they place important constraints on theoretical explanations of false memory.

The recent introduction of new techniques for studying false memories has permitted researchers to systematically identify and investigate variables that influence the likelihood of false memories. One method involves presenting meaningfully associated words (e.g., *bed*, *rest*, *awake*) to elicit false memory for a related but nonpresented word (*sleep*). The ability of such lists to elicit false recall was first noted by Deese (1959) and was replicated and extended to false recognition by Roediger and McDermott (1995). The method of using converging semantic associates to induce false memories has sometimes been referred to as the DRM (Deese/Roediger–McDermott) paradigm. Phonological associates (e.g., *sweep*, *steep*, *sleet*) can also induce similar false memory effects (e.g., *sleep*). Specifically, false memories can be elicited by presenting lists of phonologically related words in both recognition (Anisfeld, 1969; Schacter, Verfaellie, & Anes, 1997; Watson, Balota, & Roediger, 2003) and recall (McDermott & Watson, 2001; Sommers & Lewis, 1999; Watson et al., 2003) tasks.

The influence of many independent and subject variables on the likelihood of false memories in the DRM paradigm has been examined. One finding is that the orienting task invoked at encoding exerts an influence on the probability of false recall and false recognition (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001; Toglia

et al., 1999). Specifically, deeper, semantic encoding leads to higher probabilities of false memory (and veridical memory) than does superficial encoding: false memories for the nonpresented associates show patterns that are similar to what is known as the level of processing effect (Craik & Lockhart, 1972).

The source of this empirical finding, however, is not well defined, in part because the use of only semantically associated lists limits interpretive power. Does semantic processing enhance the likelihood of false memories across the board? Or is semantic processing especially influential when the stimuli are semantic associates, so that false memories induced by the encoding of phonological associates might be more frequent following phonological processing? The present experiments were designed to address these questions.

Results from several earlier studies provide tentative empirical grounds for the prediction that false memories arising from phonologically associated lists may indeed be enhanced by phonological encoding in comparison with semantic encoding. The false recognition probabilities for phonological associates from two such studies did not differ as a function of encoding task (semantic or phonological) (Coltheart, 1977; Wright, Ciccone, & Brelsford, 1977). However, two other studies noted significantly greater probabilities of false recognition for phonological encoding (relative to semantic encoding), although the observed differences were small (Davies & Cubbage, 1976; Parkin, 1983). Further interpretive complications for these studies arise because a floor effect occurred for false memories when base rate false alarms were accounted for (i.e., false recognition probabilities were in the range of 1%–5%). Nonetheless, integration of the recent associative false memory literature and the results of these studies provide a tentative empirical basis for the

prediction that a crossover interaction might occur when orienting task (semantic, phonological) is crossed with associative list structure (semantic, phonological). That is, attention to semantic relations, as compared with attention to phonological relations, may increase the likelihood of false memories for lists of semantically associated words. Conversely, attention to phonological relations, in comparison with attention to semantic relations, may increase the likelihood of false memories for lists of phonologically associated words.

Not only are there empirical grounds for predicting the crossover effect, but recent theoretical frameworks for understanding false memories lead to the same prediction. To the extent that false memories arise at least in part from associative activation (as suggested by McDermott & Watson, 2001; Roediger et al., 2001), attending to semantic features will selectively enhance the likelihood of false recall and false recognition for semantic associates (but not for phonological associates) relative to attending to phonological features. This prediction arises from the literature showing that the amount of associative priming can be mediated by attention (Balota, Black, & Cheney, 1992); priming of *mountain*, for example, should be enhanced by attention to relations among its associates (*hill*, *valley*). This enhanced priming will—other things being equal—result in enhanced probabilities of false recall and false recognition. Similarly, one would expect that attending to phonological (relative to semantic) relations among phonologically associated words would enhance activation of nonpresented phonological associates, thereby enhancing later false memory for these nonpresented associates.

A similar prediction arises from a consideration of the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), which posits that to the extent that overlap exists in features for overt and covert events, recollection for what occurred becomes confusable with what was only thought about, inferred, or activated during encoding. The present experiments do not differentiate between these two frameworks. Rather, they test a prediction made clearly by these theories against an alternative possibility—that semantic processing may enhance the likelihood of false recall and false recognition (relative to phonological processing), regardless of the to-be-remembered verbal materials.

In Experiment 1, the probabilities of accurate and false recall were measured after participants studied lists of 16 phonological associates. Encoding strategies were manipulated via orienting instructions. Specifically, participants were asked to process the words within each list with a *meaning* or a *sound* orienting task (i.e., to think about the semantic relations or the phonological relations among words in the list). We found (to anticipate the results) that the phonological orienting task did lead to greater probabilities of false recall than did the semantic orienting task. To allow for a direct comparison between results obtained for lists of semantic and phonological associates (and observation of a materials × encoding strategy interaction),

we included the semantic lists in addition to the phonological lists in a free-choice recognition test in Experiment 2.

EXPERIMENT 1

Method

Participants. The participants were 39 undergraduate students at Washington University. They were paid $5 for their participation.

Design and Materials. The lists of phonological associates are reported in the appendix of Watson et al. (2003; these lists were also used by McDermott & Watson, 2001). Changes were made to two items in the *glass* list because they were nonwords; this change was necessary given that we used an orienting task in which word meanings were analyzed. The replacement words were chosen using the same guidelines used in creating the original lists. A pilot experiment (not reported) allowed us to use the 24 lists that most effectively induce false recall.

In a within-subjects design, participants studied twelve 16-word lists for the meaning orienting task and the other 12 lists for the sound orienting task. The list type–orienting task combinations were counterbalanced across participants. Study lists were presented in a different random order for each participant, but the order of words within lists, which was randomly determined, was held constant across participants.

Procedure. Participants were tested individually or in groups of two to three. They sat in front of a computer with dividers separating the computer terminals. Encoding instructions were presented on a computer monitor. The semantic orienting instructions asked the participants to concentrate on the similarities in the meanings of the words within each list, whereas the phonological orienting instructions asked them to concentrate on the similarities in the sounds of the words. To aid understanding of the meaning orienting task, the following example was provided:

> Concentrating on the meaning and sound of the words will sometimes be difficult because the words within the lists will generally be related by sound. Nonetheless, when we ask you to focus on meaning, please do so; think about how each word relates to the meaning of the other words in the list.
> For example, you might see words like *house*, *mouse*, *couch*, etc. and be asked to pay attention to their meaning. In this case you might think about how both mouse and couch can be found in a house.

A cue word (either MEANING or SOUND) informed the participants to attend to the meaning or sound relations among the words in the upcoming list. The MEANING cue was presented in red, and the SOUND cue was presented in blue. This cue word stayed on the computer screen for 2 sec. After a 500-msec blank interval, the first list word was shown. Each list word was shown for 800 msec, with a 500-msec interstimulus interval.

After the presentation of each word list, a visual cue signaled the participants that they should begin recall. They were given 60 sec to write down as many words as they could remember from the previous word list and were told not to guess. The experimenter also repeatedly emphasized that the participants should concentrate on performing the orienting task according to the cue and not worry about the implications of this task for subsequent recall performance. At the end of the 60-sec recall period, the computer screen turned red (for 5 sec) with an accompanying audio tone, at which point the participants stopped writing and prepared for the next list. Participants were also asked to cover their answers for all previous trials with a paper and not to make any revisions or changes to their answers for the earlier lists.

Results and Discussion

An inspection of Figure 1 suggests that the sound task led to a greater probability of false recall than did the

meaning task, whereas the meaning task conferred a slight advantage (relative to the sound task) for accurate recall. All analyses were done with an alpha level of .05. Partial eta squared indicates effect size.

A 2 × 2 repeated measures analysis of variance (ANOVA) for response type (accurate or false) and orienting task (meaning or sound) revealed a reliable main effect of response type [$F(1,38) = 56.93$, $MS_e = 0.02$, $\eta^2 = .60$], which indicated that accurate recall (.41) exceeded false recall (.23). The main effect of orienting task was also reliable [$F(1,38) = 17.49$, $MS_e = 0.009$, $\eta^2 = .32$]; that is, phonological processing produced higher levels of recall (.35) than did semantic processing (.29). The interaction between response type and orienting task was also reliable [$F(1,38) = 37.16$, $MS_e = 0.009$, $\eta^2 = .494$]. Specifically, the meaning orienting task produced a slightly higher probability of accurate recall (.42) than the sound orienting task (.39) [$t(38) = 2.36$, $SEM = 0.01$]. Most importantly, the reverse pattern occurred for false recall: The sound task greatly enhanced false recall relative to the meaning task [$t(38) = 5.66$, $SEM = 0.03$]. Participants were far more likely to recall the critical nonpresented word if they had encoded the word list with respect to its sound (.31) rather than its meaning (.16) characteristics. This pattern of false recall, as depicted in Figure 1, represents a phonological superiority effect that hitherto has not been observed in the DRM false memory literature. A theoretical consideration of these findings will be delayed until the General Discussion.

EXPERIMENT 2

Two goals were pursued in Experiment 2. First, we sought to replicate and extend the phonological superiority effect for false memory of phonological associates. Specifically, we sought to extend the pattern to free-choice recognition. Second, we sought to demonstrate the full crossover interaction with the addition of semantically associated lists. As mentioned in the introduction to this article, prior studies have repeatedly shown that semantic processing elicits greater false recall and false recognition for semantically associated lists. Therefore, for the results of Experiment 1 to be convincing, it is necessary to show that the phonological superiority effect in false memory occurs selectively for phonologically associated lists and that we can replicate the standard finding of semantic superiority in false memory for semantically associated lists.

Method

Participants. A total of 54 undergraduates at Washington University participated either for research credit or $5.

Design and Materials. We used the same 36 semantic and 36 phonological lists used in McDermott and Watson (2001) and Watson et al. (2003). The semantic and phonological lists converged on the same 36 critical words. The participants studied 24 lists (12 semantic, 12 phonological) during encoding: The remaining 12 lists served as base rate items, which were counterbalanced across participants. In a 2 × 2 within-subjects design, participants studied 6 of the semantic lists for a semantic processing task, 6 of the semantic lists for a phonological processing task, 6 of the phonological lists for a se-
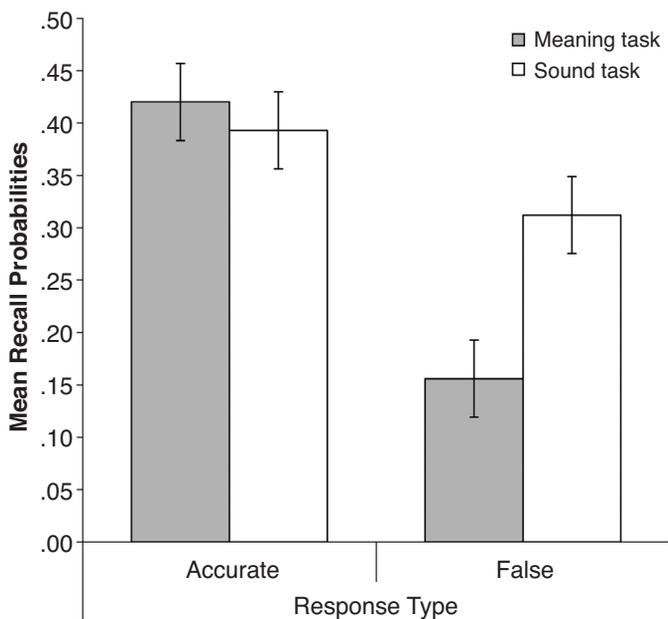

Figure 1. Mean probabilities of immediate free recall for the phonologically related words as a function of response type (accurate and false recall) and orienting task (meaning and sound) for Experiment 1. Error bars display .95 confidence interval.

mantic processing task, and 6 of the phonological lists for a phonological processing task. As in Experiment 1, the 24 lists were presented in a unique random order for each participant.

All participants received the same 108 words for the recognition test. These 108 words consisted of 36 critical lures, 36 words from the semantic lists (one word selected randomly from each list), and 36 words from the phonological lists (again, one word selected randomly from each list). Of these 108 words, 84 were used in our analysis and could be classified into the following seven groups: (1) One word from each of the 12 semantic lists presented during encoding. These words, if correctly recognized, were scored as hits separately for the meaning and sound tasks. (2) One word from each of the 12 phonological lists presented during encoding. Like the words in Group 1, these were also scored as hits separately for the two tasks. (3,4) One critical word from each of the 24 studied lists. These words, if falsely recognized, were scored as critical false alarms separately for the semantic (3) and the phonological (4) lists (and separately for the two orienting tasks). (5) One semantic associate from each of the 12 lists not presented during encoding. These words served as base rate false alarms for the semantic lists. (6) One phonological associate from the 12 lists not presented during encoding. These words served as base rate false alarms for the phonological lists. (7) One critical word from the 12 lists not presented during encoding. These words were scored as base rate false alarms for the critical words.[1]

Procedure. The stimulus presentation protocol and encoding instructions were the same as in Experiment 1, but with three exceptions. First, the participants were not given specific examples of how to process words on the incongruent dimension (i.e., sound processing for semantic lists and meaning processing for phonological lists); they were simply asked to attend to either the sound or the meaning of words. Second, encoding in this experiment was incidental; the participants were told about the recognition test only after all 24 lists had been presented. Third, the participants were asked to provide a relatedness rating immediately following the presentation of each word list on the cued dimension (semantic or phonological). The relatedness ratings were spread across a 7-point Likert scale, with 1 indicating *extremely unrelated* and 7 *extremely related*. For example, if a word list was studied with the meaning orienting task, the participants were asked to judge how related the 16 words were as a whole on the meaning dimension, regardless of whether it was a semantic or a phonological list. This relatedness judgment served as a cover task for incidental learning; therefore, results of this task were not analyzed in depth.[2]

The participants were given instructions for the free-choice recognition test after the encoding phase. They were instructed to press the *Y* key for words that they had seen during encoding, and the *N* key for words that they had not seen during encoding.

Results and Discussion

Figure 2 shows the corrected recognition rates (with base rate false alarms subtracted out) for the four encoding conditions. Consider first the data from the semantically associated lists (shown on the left side of the figure). The meaning task led to higher probabilities of accurate and false recognition than did the sound task, a pattern that conceptually replicated prior findings (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001; Toglia et al., 1999). The opposite pattern occurred for phonologically associated lists (as can be seen on the right side of the figure): The sound task elicited higher probabilities of accurate and false recognition than did the meaning task.

All of the analyses conducted and data reported for this section were based on corrected recognition rates
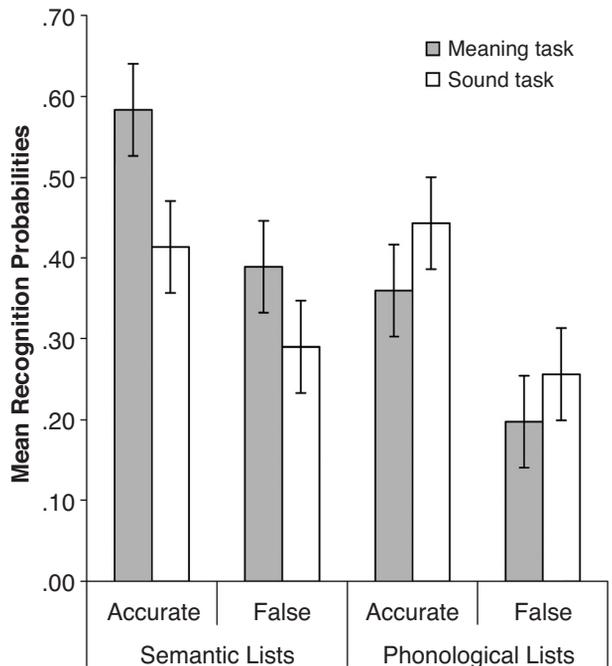

Figure 2. Mean probabilities of final free-choice recognition as a function of list type (semantic and phonological), orienting task (meaning and sound), and response type (accurate and false recognition) for Experiment 2. Error bars display .95 confidence interval.

(see Table 1 for raw recognition probabilities and their corresponding base rates).

Overall recognition probabilities. The three-way interaction for list type (semantic or phonological), response type (correct or false recognition), and orienting task (meaning or sound) was not reliable [$F(1,53) = 1.9$, $p = .18$]. In addition, orienting task and response type did not interact reliably ($F < 1$). Importantly, as predicted, list type and orienting task produced a crossover interaction [$F(1,53) = 43.69$, $MS_e = 0.03$, $\eta^2 = .452$]. This interaction arose because the same orienting task led to opposite patterns of results depending on list type (see Figure 2).[3] We examine the data from the semantic and phonological lists separately in the following analyses.

Table 1
Mean Recognition Probabilities of Experiment 2

List Type	Task	Response Type		Base Rate (Hits)	Base Rate (FA)
		Hits	FA		
Semantic	Meaning	.76 (1)	.64 (3)	.18 (5)	.25 (7)
	Sound	.59 (1)	.54 (3)		
Phonological	Meaning	.61 (2)	.45 (4)	.25 (6)	
	Sound	.69 (2)	.51 (4)		

Note—Numbers in parentheses refer to the seven item groups described in the Method section. Hits, hit rate for the studied list associates; FA, false alarm rate for the critical nonpresented words; Base Rate (Hits), base rate false alarms for the studied words; Base Rate (FA), base rate false alarms for the critical nonpresented words.

Semantic lists. No significant interaction was found between orienting task and response type [$F(1,53) = 1.6$, $p = .21$]; however, the main effects for both orienting task [$F(1,53) = 29.6$, $MS_e = 0.03$, $\eta^2 = .358$] and response type [$F(1,53) = 15.3$, $MS_e = 0.09$, $\eta^2 = .224$] were significant. The main effect of orienting task indicated that meaning processing (.49) led to higher levels of overall recognition than did sound processing (.35). The main effect of response type showed that participants were more likely to recognize studied words (.50) than critical nonpresented words (.34). The lack of an interaction between orienting task and response type indicated that accurate and false recognition behaved similarly under the orienting task manipulation. Planned comparisons revealed that the meaning task produced a higher level of hits (.58) than the sound task (.41) [$t(54) = 4.66$, $SEM = 0.036$]. A similar effect was observed for false alarm rates—the meaning task led to a higher level of false alarms (.39) than the sound task (.29) [$t(54) = 2.58$, $SEM = 0.038$].

Phonological lists. Significant main effects were found for orienting task [$F(1,53) = 9.41$, $MS_e = 0.03$, $\eta^2 = .151$] and response type [$F(1,53) = 20.94$, $MS_e = 0.08$, $\eta^2 = .283$], but these two variables did not interact ($F < 1$). The main effect of response type indicated that studied words were called old (.40) more often than critical nonpresented words (.23), whereas the main effect of orienting task arose because the sound orienting task led to a higher recognition probability (.35) than the meaning orienting task (.28). Planned comparisons showed that the sound task produced a higher probability of hits (.44) than the meaning task (.36) [$t(54) = 2.13$, $SEM = 0.039$]. A parallel pattern was observed for false alarms: The sound task led to a higher probability of false alarms (.26) than the meaning orienting task (.20) [$t(54) = 2.03$, $SEM = 0.029$]. To our knowledge, this is the first empirical demonstration of a reverse levels-of-processing effect in a standard recognition test. This finding is especially noteworthy in that it differs from other reverse levels-of-processing findings (Morris, Bransford, & Franks, 1977) in which memory performance from two (or more) tests with explicitly different retrieval demands was compared.

GENERAL DISCUSSION

Three key findings emerged from these two experiments. First, when lists of converging phonological associates were encoded and a free recall test followed, the likelihood of false recall of a related, nonpresented associate was greater when the orienting task required processing the phonological characteristics of the words than when it required processing their semantic characteristics (Experiment 1). Second, this pattern was extended to false recognition: Processing sound features of a list of phonological associates led to greater false recognition of related, nonpresented associates than did processing meaning features (Experiment 2). Experiment 2 further demonstrated that this effect occurred in a situation in which the opposite pattern was observed for semantic lists, which replicated findings from prior studies. That is, when lists of converging semantic associates were encoded, an orienting task that focused on the meaning of words enhanced false recognition relative to an orienting task that focused on the sound of words. Third, when phonologically associated lists were presented, sound-oriented processing enhanced later accurate recognition (relative to meaning-oriented processing) on a free-choice recognition test. We now consider these findings from a broader perspective.

The present findings are consistent with extant theories of false memory. Consider the activation/monitoring framework, as outlined by McDermott and Watson (2001) and Roediger et al. (2001). This framework draws in part on the idea of spreading activation (Collins & Loftus, 1975), which holds that related concepts are linked together in semantic memory via a series of pathways. Accessing one concept (e.g., *hill*) sends activity across the pathways to related concepts (e.g., *valley*, *mountain*), thereby making these related concepts more accessible (i.e., primed). An important consideration is that the amount of associative priming achieved can be modulated by strategic processes (Balota et al., 1992; Neely, 1977). Such modulation could happen either via inhibition (when the orienting task and materials mismatch) or augmentation (when the orienting task and materials match).

Thus, the activation/monitoring approach predicts that the priming achieved via presentation of associative lists could be enhanced by appropriate strategic processes and that this enhanced priming would—other things being equal—translate into an enhanced likelihood of false memories. Specifically, a match between material and processing (i.e., attention to phonological relations among phonologically associated words or to meaning relations among semantically associated words) should boost activation of the critical nonpresented word, thus enhancing the possibility of its false recall or false recognition. The results presented here are fully consistent with this viewpoint. Note that this theoretical framework is not the only one that can accommodate the present results. Rather, we point out how readily this framework makes a clear prediction that is upheld in the present report. Similar predictions would arise from (e.g.) the source monitoring framework (Johnson et al., 1993) and fuzzy trace theory (Reyna & Brainerd, 1995).[4]

To our knowledge, the finding that phonological processing, relative to semantic processing, can enhance retrieval of studied words on a standard free-choice recognition test is a new one (cf. Intraub & Nicklos, 1985, who obtained a similarly puzzling effect for pictures.). Although there have been numerous studies in the transfer-appropriate processing and encoding-specificity literatures showing that in certain situations phonological or orthographic orienting tasks can boost memory performance relative to semantic orienting tasks, these studies have used nonstandard test conditions (e.g., Morris et al.,

1977). For example, a rhyme recognition test requires that subjects recognize whether a target word presented at test rhymes with a studied word. The present results show that with a powerful manipulation such levels-of-processing effects can be reversed even on a standard free-choice (or yes/no) recognition test. However, along with this unanticipated, interesting finding comes an unanticipated puzzle in that the free recall test (Experiment 1) did not show a reverse levels-of-processing effect. To appreciate this cross-test difference, consider the two leftmost bars in Figure 1 and the 5th and 6th bars (from the left) in Figure 2. That is, for phonological lists, the probability of accurate recall was greater for the meaning orienting task (.42) than for the sound orienting task (.39). However, for the same lists, the probability of accurate recognition was greater for the sound task (.44) than for the meaning task (.36). We consider here some thoughts regarding potential explanations for this difference.

One possible framework for thinking about this difference (and for thinking about our experiments on the whole) is the material-appropriate processing framework (Einstein, McDaniel, Owen, & Coté, 1990; McDaniel & Einstein, 1989). This framework holds that accurate recall will benefit to the extent that the prior encoding task encourages processing of attributes that are not invited naturally by the stimuli. For example, a list of rhyming words should naturally foster phonological processing; therefore, asking participants to perform an orienting task that attends to the phonological relations among these words would be redundant and ineffective. In contrast, asking participants to perform an encoding task that attends to the nonobvious attributes of the stimuli would provide extra item-specific processing on top of the relational processing already encouraged by the stimuli, thus leading to superior recall performance. To put this idea into the present context, a semantic task performed on a list of phonological associates and a phonological task performed on a list of semantic associates would be expected to enhance performance for studied words (relative to the conditions in which the type of processing afforded by the stimulus and that performed by the subject matched).

The prediction of superior accurate recall performance following a material-appropriate (essentially a material–encoding mismatch) task, relative to a material–encoding match task, was borne out in the accurate recall data from Experiment 1; that is, for the phonological associates, semantic encoding led to better accurate recall than did phonological encoding. This finding, however, was reversed in Experiment 2. To be fair, the material-appropriate processing framework was developed entirely under recall experiments; therefore, it is unclear whether the framework was intended to apply to recognition in the same manner. We have considered this possible explanation of our accurate recall and recognition data because we find the pattern obtained in the recognition test to be highly interesting and to possibly provide fertile ground for future research. We note, however, that the cross-experiment differences require further empirical findings to reach any firm conclusions.

In summary, the present experiments have demonstrated that interactions between encoding task and materials can be crucial determinants of later false recall. Specifically, attention to associations along the dimensions in which the concepts are related fosters later false memory.

REFERENCES

ANISFELD, M. (1969). False recognition produced by semantic and phonetic relations under two presentation rates. *Psychonomic Science*, **17**, 366-367.

BALOTA, D. A., BLACK, S. R., & CHENEY, M. (1992). Automatic and attentional priming in young and older adults: Reevaluation of the two-process model. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 485-502.

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

COLTHEART, V. (1977). Recognition errors after incidental learning as a function of different levels of processing. *Journal of Experimental Psychology: Human Learning & Memory*, **3**, 437-444.

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.

DAVIES, G. M., & CUBBAGE, A. (1976). Attribute coding at different levels of processing. *Quarterly Journal of Experimental Psychology*, **28**, 653-660.

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

EINSTEIN, G. O., McDANIEL, M. A., OWEN, P. D., & COTÉ, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory & Language*, **29**, 566-581.

INTRAUB, H., & NICKLOS, S. (1985). Levels of processing and picture memory: The physical superiority effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 284-298.

JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.

McDANIEL, M. A., & EINSTEIN, G. O. (1989). Material-appropriate processing: A contextualist approach to reading and studying strategies. *Educational Psychology Review*, **1**, 113-145.

McDERMOTT, K. B., & WATSON, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory & Language*, **45**, 160-176.

MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.

NEELY, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**, 226-254.

PARKIN, A. J. (1983). The relationship between orienting tasks and the structure of memory traces: Evidence from false recognition. *British Journal of Psychology*, **74**, 61-69.

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

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

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.

ROEDIGER, H. L., III, WATSON, J. M., McDERMOTT, K. B., & GALLO, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, **8**, 385-407.

SCHACTER, D. L., VERFAELLIE, M., & ANES, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology*, **11**, 331-342.

SOMMERS, M. S., & LEWIS, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory & Language*, **40**, 83-108.

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

TOGLIA, M. P., NEUSCHATZ, J. S., & GOODWIN, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, **7**, 233-256.

WATSON, J. M., BALOTA, D. A., & ROEDIGER, H. L., III (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory & Language*, **49**, 95-118.

WRIGHT, J., CICCONE, D. S., & BRELSFORD, J. W. (1977). Selective encoding processes in recognition. *British Journal of Psychology*, **68**, 289-295.

NOTES

1. Two other groups of words were included in the recognition test but were excluded from the analyses. We did not use these as our base rate items because of possible contamination. The first group consisted of the phonological associates of the 12 semantic lists presented during encoding. For example, if participants have studied words such as *bed*, *rest*, *awake* and implicitly activated the word *sleep*, the erroneous recognition of an item such as *steep* might not reflect the true base rate false alarm for these words, because the implicit activation of the word *sleep* might consequently have activated the word *steep* (an item from the phonological version of the *sleep* list) through its phonological neighborhood. The second group included the semantic associates of the 12 phonological lists presented during encoding. Specifically, if participants studied words such as *steep*, *weep*, *sheep* and implicitly activated the word *sleep*, the erroneous recognition of an item such as *bed* might not reflect true base rate false alarm, because the implicit activation of the word *sleep* might consequently have activated the word *bed* by means of spreading activation.

2. The average relatedness ratings for the four list–task combinations are as follows (standard deviations are presented in parentheses): 6.01 (.91) for semantic list–meaning task, 1.43 (.57) for semantic list–sound task, 1.68 (.78) for phonological list–meaning task, and 5.42 (.81) for phonological list–sound task.

3. We also conducted all analyses using the two groups of items mentioned in Note 1 as our base rate false alarms. The corrected hit rates using these base rate items were .61 for semantic list–meaning task, .44 for semantic list–sound task, .33 for phonological list–meaning task, and .41 for phonological list–sound task. As can be seen, these data lead to the same conclusions as our primary analyses.

4. Strategic processes such as source monitoring also play an important role during the retrieval phase when subjects have to assess whether highly activated concepts were studied or simply primed; this retrieval aspect of monitoring has been discussed in the introduction to this article. As we have noted earlier, our experiments were not designed to distinguish between activation- and monitoring-related frameworks; although such distinction is important, it is not the primary focus of the present discussion.