Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Journal of Memory and Language 52 (2005) 1–28

Journal of Memory and Language

www.elsevier.com/locate/jml

Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures[☆]

Keith A. Hutchison[a,*], David A. Balota[b]

[a] *Department of Psychology, 304 Traphagen Hall, Montana State University, Bozeman, MT 59717-3440, United States*
[b] *Washington University in St. Louis, United States*

Received 12 May 2004; revision received 7 August 2004

Abstract

Veridical and false memory were examined in lists that contained 12 words that all converged onto the same meaning of a critical nonpresented word (e.g., snooze, wake, bedroom, slumber. . . , for SLEEP) or lists that contained 6 words that converged onto one meaning and 6 words that converged onto a different meaning of a homograph (e.g., stumble, season, trip, autumn. . . , for FALL). Associative strength from the list items to the critical item was equated across the two types of lists. In Experiments 1–5, patterns of veridical memory differed across the two types of lists; however, false memory of the critical item did not differ. This same pattern occurred regardless of whether the words diverging onto the two meanings of the homograph were presented blocked or intermixed, whether each list item was presented for 80 ms, 200 or 1200 ms during encoding, and whether a recall or recognition test was given. In Experiment 6, critical nonpresented items that followed lists converging onto one meaning were judged as more strongly related to the list. These results suggest that false memory in the DRM paradigm largely reflects lexical/associative activation, rather than the formation of a meaningful thematic representation.

Keywords: False memory; Recall; Recognition; Semantic; Associative; Homograph; Priming

Aristotle speculated that humans organize the world into a coherent mental representation through the formation of links between related experiences. In particular, his laws of association governed that such links are likely to be formed between concepts that are similar, opposites, or follow one another closely in time. More recently, cognitive psychologists have attempted to formalize a mechanism through which such extensive associative networks could (1) represent a vast amount of world knowledge and (2) access such knowledge to answer general knowledge questions, predict upcoming events, and make inferences during comprehension

☆ Some of the experiments reported were conducted while the first author was a post-doctoral research associate at Washington University in Saint Louis. The authors would like to thank Doug Nelson and Henry L. Roediger for helpful comments on an earlier version of this manuscript. The authors also thank Jesse Bengson, Justin Deem, Jonathan Ganger, Virginia Johnson, Jeff Templeton, Marlena Wu, and Josie Zelzer for their help in constructing lists, running participants, and data entry.
* Corresponding author. Fax: +1 406 994 3804.
E-mail address: khutch@montana.edu (K.A. Hutchison).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2004.08.003

(Anderson, 1983; Collins & Loftus, 1975; Kintsch, 1974; Ratcliff & McKoon, 1978).

The most common procedure for investigating the organization of such networks of related information is the *semantic priming* paradigm (Anderson, 1983). Using this paradigm, researchers have discovered that responding to a target word such as "cat" is faster (in naming and lexical decision tasks) following a semantically related prime (e.g., dog) than following an unrelated prime (e.g., table). Because relatedness exerts an influence in these simple tasks, some researchers have suggested that semantic priming reflects an *automatic spreading activation* mechanism in which, while reading or hearing a word, activation automatically spreads from the semantic representation (node) of that word to the representations (nodes) of semantically associated neighbors (Neely, 1977; Posner & Snyder, 1975).

Demonstrations of semantic relatedness have also been obtained using episodic memory tasks. For instance, Underwood (1965) noticed that the presence of a word such as "table" in a study list increased people's likelihood of falsely recognizing a related word such as "chair" during a later recognition test, relative to an unrelated word such as "screen." In a more powerful procedure, known as the Deese–Roediger–McDermott (DRM) false memory paradigm (after Deese, 1959; Roediger & McDermott, 1995), participants see or hear lists that include the first 15 associates for a given target word and are then given a recall or recognition test. The robust finding from these studies is that the nonpresented target word is falsely remembered at very high levels (see Gallo & Roediger, 2002; Roediger, Balota, & Watson, 2001; Roediger & Gallo, 2003; for recent reviews). In fact, in some circumstances, these words are recalled or recognized as often (Roediger & McDermott, 1995) or even more often (Brainerd & Reyna, 1998; McDermott, 1996; Watson, Balota, & Roediger III, 2003) than items actually presented.

The question of meaning

For both semantic priming and false memory paradigms, one fundamental question has centered on whether the effect reflects lexical associative activation from the prime (or studied items) to the target (or critical nonpresented item) or is due to the extraction of meaning from the prime which then facilitates the processing of the target. The difficulty in answering this question stems from the fact that both priming studies and false memory studies rely heavily on stimuli obtained from word association norms. A vast majority of such associated pairs contain a large overlap in semantic features (see Table 1 from Hutchison, 2003). For instance, the words "cat" and "dog" are both associated (in that they typically co-occur in language) and semantically related (in that they are part of the same PET category and share many semantic features such as "fur" and "claws"). As a result, priming effects from such items could be due to either lexical association, semantic feature overlap, or both.

We will briefly review the evidence from automatic semantic priming tasks and argue that priming may simply be due to associative activation, rather than semantic feature overlap. We will then describe a paradigm from "semantic" priming studies which affords a way to discriminate associative vs. semantic priming. Finally, we will explain how this technique can be implemented to explore the role of meaning on false memories in the DRM paradigm.

Association vs. meaning in semantic priming

Fodor (1983) proposed simple associative links between words that tend to co-occur either in experience or in language so that "co-occurrence relations among *mental* events mirror the corresponding relations among *environmental* ones" p. 33. Accordingly, reading or hearing the word "salt" will automatically activate "pepper," not because these words have similar meanings, but because they tend to co-occur together. Researchers have more recently argued that semantic priming is not due simply to spreading activation across such associative links, but rather to shared semantic features (e.g., salt and pepper are both small, both spices, both found in shakers, etc...) between primes and targets (Kawamoto, 1993; Masson, 1995; Moss, Hare, Day, & Tyler, 1994; Plaut, 1995). Recent priming publications appear to support this assumption (see Lucas, 2000; for a review).

Hutchison (2003) has recently reviewed the studies used as support for the feature overlap hypothesis and came to a different conclusion. Specifically, he argued that there was no strong evidence of automatic priming for items lacking an association (e.g., "horse-deer," see Lupker, 1984; Shelton & Martin, 1992; for similar conclusions). For example, when re-examining the stimuli used to support the importance of feature overlap to priming (e.g., de Morney Davies, 1998; Hines, Czerwinski, Sawyer, & Dwyer, 1986; Thompson-Schill, Kurtz, & Gabrieli, 1998), Hutchison found that the "semantic" items were more strongly associated than the "associative" items, according to the Nelson, McEvoy, and Schreiber (1999) word-association norms. Thus, the larger priming for "semantic" items could have just as easily reflected association strength as featural overlap. In contrast, Hutchison found strong support for automatic priming based only on association. For example, items that share little or no features (e.g., LION-STRIPES) but are associated via a "mediating" associative link (e.g., TIGER) show consistent priming effects, and such effects occur under relatively automatic priming conditions (e.g., Balota & Lorch, 1986; McNamara & Altarriba, 1988).

Balota and Paul (1996) provided additional evidence for the importance of associative links in priming. These authors tested the lexical associative vs. semantic feature hypotheses using multiple prime words that were either associated to two different meanings of the same homographic target (e.g., kidney-piano-ORGAN) or the same meaning of a nonhomographic target (e.g., lion-stripes-TIGER). These different levels of representation are shown in Fig. 1. To test the influence of the primes, they used four conditions for each stimulus type: An unrelated (UU) condition (e.g., wagon-soda-organ), a RU condition (e.g., kidney-soda-organ), an UR condition (e.g., wagon-piano-organ), and a RR condition (kidney-piano-organ). Priming effects for each condition were obtained by subtracting reaction-times in each related condition from the unrelated (UU) condition.

If priming were due to lexical activation, priming should be similar for homographic and nonhomographic targets, since both prime words are associated to the same lexical entry. Alternatively, if priming were due to the semantic overlap between primes and targets, there should be relatively reduced priming for the homographic target compared to the nonhomographic target, since for homographs the two primes diverge onto two unrelated meanings (e.g., the body and musical instrument interpretations of ORGAN). Across 5 experiments, priming effects were remarkably similar for homographic and nonhomographic targets.

Thus, multiple associative primes produced additive effects when combined, and these additive effects emerged regardless of whether the primes converged upon the same conceptual meaning of the target or instead diverged unto two distinct conceptual meanings of the target. As argued by Balota and Paul, this pattern is consistent with the lexical association hypothesis that priming is determined by the lexical associations between primes and targets, but is inconsistent with the featural overlap hypothesis that priming reflects the conceptual overlap between primes and targets.

To show that semantic overlap can play a role under conditions in which attention to semantic-based representations is required, Balota and Paul (1996) conducted a further experiment in which participants performed a relatedness judgment on the target (“was one or both of the primes related to the target?”). Under these conditions, they indeed found a difference between the two types of stimuli, with an additive effect for nonambiguous targets (i.e., RR priming = RU priming + UR prim-

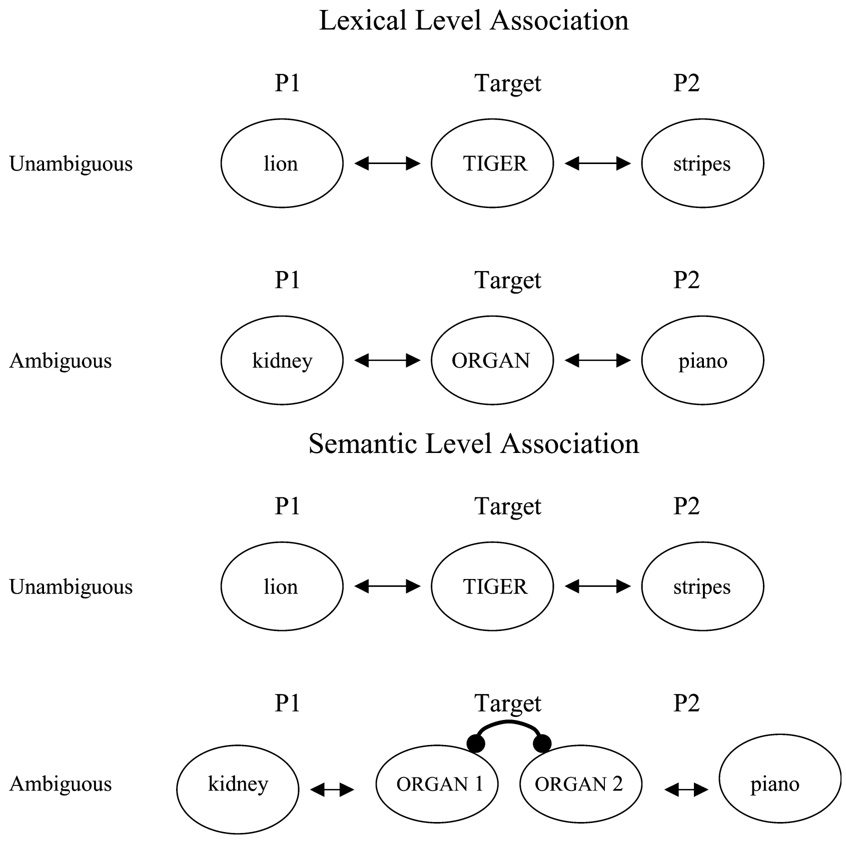

Fig. 1. Example of lexical and semantic-level associations (figure taken from Balota & Paul, 1996).

ing) and an under-additive effect for ambiguous targets (i.e., RR priming < RU priming + UR priming). Balota and Paul concluded that additional effects of "meaning" (above those based purely from association), require the conscious direction of attention to semantic-based representations, such as in relatedness decisions. Hence, these data are consistent with the conclusions of Hutchison (2003) that semantic priming may simply reflect an automatic associative activation process. Any further influence of conceptual representations might occur consciously if such information is relevant to the decision process.

Meaning in episodic memory

Although conceptual representations may play a relatively small role in semantic priming, they undoubtedly play a stronger role in tasks of episodic memory. In fact, the importance of attending to an item's meaning during study is one of the earliest and most well-documented findings in the memory literature (see Crowder, 1976; for a review). For instance, Bousfield (1953) observed that words from the same category tended to be "clustered" together during free recall, even though all items were presented in random order during study. Later, Thompson, Hamlin, and Roenker (1972) found a positive relation between participants' amount of clustering and their total recall performance. Miller (1956) argued that combining list words into categorized "chunks" allows participants to bypass the limited capacity of immediate memory, because the category labels can help generate the studied exemplars during retrieval. Cohen (1963) gave participants either lists consisting of between 10 and 20 categories (with 3–4 items in each category) or lists consisting of between 10 and 20 unrelated words and found that the number of categories recalled approximated the number of unrelated words recalled. Moreover, participants recalled approximately 60–80% of the items within each category. Although some authors (Deese, 1959; Jenkins & Russell, 1952) pointed out that clustering could merely reflect inter-item associations, Marshall (1967) found that category membership influences clustering above and beyond the effects of association. In fact, hundreds of papers have since supported the benefits of semantic processing on later recall, and this basic benefit underlies the well-cited levels of processing account of memory (see Craik & Lockhart, 1972).

Although semantic organization generally improves memory, it could potentially lead people to produce thematically consistent items that were never actually studied. There is considerable research showing that, when reading text, people quickly extract the overall meaning while discarding particular sentence structures that conveyed the information (Bransford & Franks, 1971). Such meaning extraction could easily produce false memories of "theme consistent" items. In the DRM paradigm, several theories of false memory have been proposed in which intrusions arise supposedly because they are consistent with the underlying meaning or theme of studied material (Arndt & Hirshman, 1998; Brainerd & Reyna, 2002; Toglia, Neuschatz, & Goodwin, 1999; Whittlesea, 2002). For instance, According to Fuzzy Trace Theory (Brainerd & Reyna, 1990), a studied list item such as "rest" would contain both a verbatim representation (e.g., visual or auditory presentation details) and a gist representation (e.g., the meaning of "rest"). Veridical memory for studied items could occur either by the retrieval of the verbatim or gist representations. In contrast, a critical nonpresented item should only contain a gist representation (e.g., the meaning of "sleep"). False memory for these items could only occur due to the retrieval of a gist representation (e.g., the meaning of list items such as "rest" or the overall "sleep" theme creating by the list items). Brainerd and Reyna (1998) point out that the individual list items in the DRM paradigm all cue similar meanings, resulting in a strong gist representation of list themes. Thus, the critical nonpresented item "sleep" is more similar to the overall gist theme than any of the actually studied items. As a result, memory for nonpresented critical items can be equal to (or even greater than) memory for studied items under conditions that favor the reliance on gist memory.

Is meaning necessary for false memories in DRM?

The thematic explanation of false memory in the DRM paradigm is qualitatively different from that originally proposed by Deese (1959) and Roediger and McDermott (1995). These authors suggested that false memories may occur due to associative activation. Namely, as intended by the construction of DRM lists, the presentation of the list items strongly activates the critical item during study. The subject then misattributes this strong activation to the item having been previously studied. In support of this argument, Deese (1959) found that total backward associative strength (BAS) from the list items to the critical item correlated at .87 with the probability of a nonpresented critical item intruding during recall. Roediger, Watson, McDermott, and Gallo (2001) later replicated the importance of BAS in a regression study comparing false memory across 55 different word lists using 7 predictor variables (length, log frequency, concreteness, FAS, BAS, connectivity, and veridical recall). BAS was by far the best predictor of false memory. Most recently, Hicks and Hancock (2002) found that BAS also influenced peoples' tendency to misattribute false memories to a particular source, with more false memories attributed to the source that presented the half of the list containing the higher BAS.

Balota et al. (1999) and Roediger, Balota et al. (2001) have emphasized the importance of associative activation, stating that false memories are created by the same automatic spreading activation process that produces speeded responding in lexical decision or pronunciation tasks. Moreover, they drew analogy to the Balota and Paul (1996) results, suggesting a summation of activation with each additional related item. An experiment by Robinson and Roediger (1997, Experiment 2), indeed suggests that the same spreading activation that produced additive priming in Balota and Paul may also produce false memories in the DRM paradigm. In their experiment, Robinson and Roediger presented participants with 15-item lists that contained 0, 3, 6, 9, 12, or 15 associates (with total BAS's of 0, .87, 1.44, 1.89, 2.28, 2.70, respectively). They found that false recall increased from .03 when only 3 items had been studied to .30 when 15 items had been studied. Moreover, although the number of intrusions increased dramatically from the 3-item list (BAS = .87, intrusions = 3%) to the 6-item list (BAS = 1.44, intrusions = 15%), the number of intrusions appeared to increase linearly with BAS thereafter. These data further support association strength as a critical determinant of false memory even when the critical item itself is held constant.

Same ole problem

Although they seem very different, testing between the meaning and activation accounts of false memory using the standard DRM procedure is extremely difficult. The main reason for this is again due to the confound in word association norms between semantic overlap and association strength that has plagued research on semantic priming. In the DRM paradigm, a critical nonpresented item could be falsely remembered either because of activation from associated items or instead because it's meaning is similar to the items that were actually studied. According to theme-based accounts of false memory, increasing thematic organization of list items should increase both veridical and false memory, with perhaps a larger increase in false memories (which are strictly gist-based) than veridical memories (which are also influenced by verbatim details). In contrast, the associative activation account predicts a dissociation, with thematic organization increasing veridical memory but having no effect on false memory.

Current study

The current series of experiments were designed to apply the Balota and Paul (1996) methodology to examine whether false memories in the DRM procedure are primarily due to lexical association or instead to semantic overlap. As with Balota and Paul, the present study used two types of items equated in associative strength to the target. For the nonhomographic lists, we used two sets of 6 associates that were all related to the same meaning of a critical target (e.g., "*slumber*, *lay*, *motel*, *trance*, *lazy*, *nightmare*" and "*snooze*, *wake*, *bedroom*, *unconscious*, *deep*, *blanket*" for the critical item SLEEP). These lists are most similar to those used in previous DRM experiments (see Stadler, Roediger, & McDermott, 1999) and are hence called the "DRM" lists. We constructed the other half of the lists such that the two sets of 6 associates diverged onto two separate meanings of the same homograph (e.g., "*stumble*, *slip*, *rise*, *trip*, *faint*, *clumsy*" and "*autumn*, *season*, *spring*, *leaves*, *brisk*, *harvest*" for the critical item FALL).

As shown in Fig. 2, this study is a conceptual replication of Balota and Paul (1996) using 12 associates rather than 2 and using free recall rather than lexical decision or pronunciation. As with Balota and Paul, we are using different sets of items across the same vs. different meaning conditions that are equated in associative strength to their respective targets (i.e., critical nonpresented items). Moreover, as shown below, we were successful in equating false recall in the six-item DRM and Homograph sets across Experiments. Of course, the critical issue is the relative *increase* in false memories when 6 additional items of either the same or different meaning are studied. The predictions of this study are straightforward. If false memories are governed by meaning, increasing the number of related items from 6 to 12 should increase both veridical and false recall for DRM lists, since the words converge on a consistent theme. However, no such increase in veridical or false memory would be predicted for homograph lists, since the additional items would be of a different meaning. If instead false memories are governed by associative activation, they should reflect only BAS from the list items. In this case, both DRM and homographic critical items should show an equal increase in false recall with an additional 6 studied associates.

Four recall experiments and 2 recognition experiments were designed to test whether false memories behave more like veridical memories (i.e., primarily influenced by meaning) or more like semantic priming (i.e., primarily influenced by associative activation). Table 1 provides an example of the organization of lists across Experiments 1–6. In Experiment 1, all words were presented for 1.2 s during study and two 12-item lists were presented at a time in a blocked manner. In Experiment 2, the two sets of associates in the 12-item lists were intermixed during presentation to investigate the influence of disrupting the thematic content for the homograph items. Experiments 3a and 3b used the same list structure as Experiment 2, but decreased the presentation duration to either 200 ms (Experiment 3a) or 80 ms (Experiment 3b) to minimize conscious strategic processing of the list items during study.

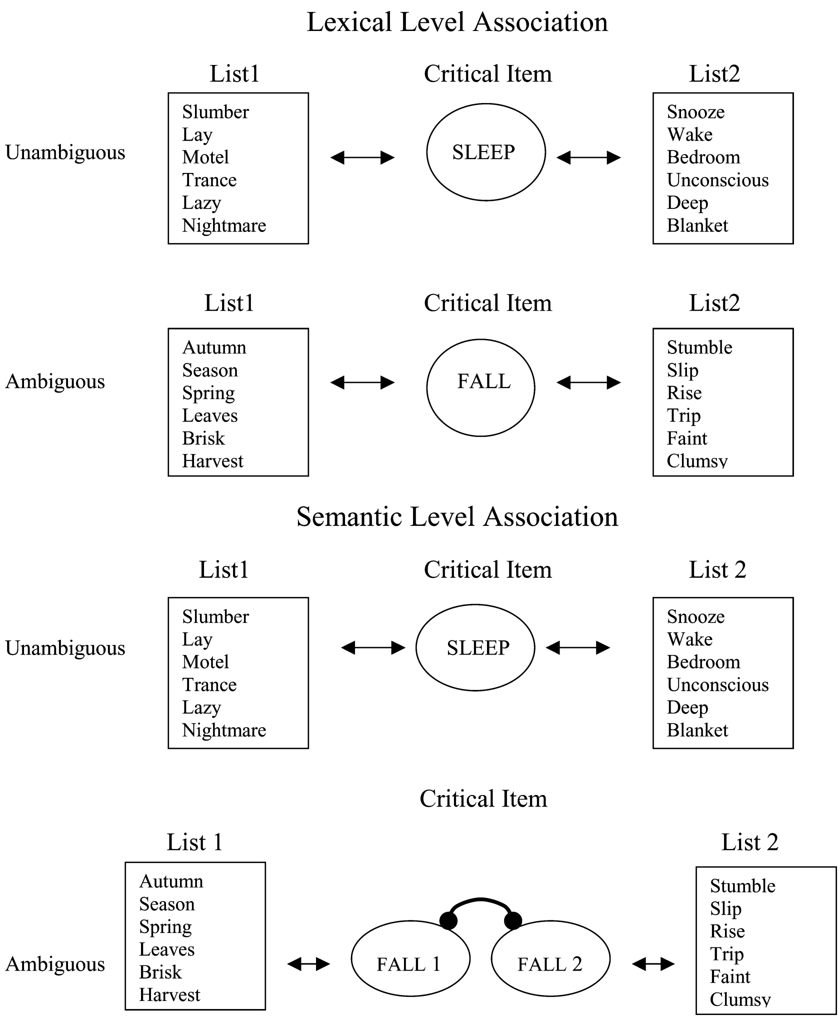

Fig. 2. Example of lexical and semantic-level associations in the current study.

Experiment 4 again used the 1.2 s presentation duration but presented only one 12-item list at a time and eliminated buffer items. In Experiment 5, participants were given a recognition, rather than recall test, after every list. Finally, in Experiment 6, the recognition decision was switched to a ‘‘relatedness’’ decision in which participants estimated each test items' degree of relatedness to the list items.

Experiment 1

Method

Participants

Thirty-six undergraduates at Washington University in Saint Louis participated for partial completion of a research requirement for an introductory psychology class. All were native English speakers with normal or corrected-to-normal vision.

Stimuli

A total of 36 sets of items were constructed for use in the experiment: Eighteen sets contained 12 words associated to a homographic item and 18 sets contained 12 words associated to a nonhomographic item.

Homograph item-set construction. Eighteen critical homograph items were taken from the Twilley, Dixon, Taylor, and Clark (1994) norms. We selected only items containing meanings that were relatively balanced, operationally defined as our ability to obtain 6 associates with some degree of BAS to the critical homograph for two separate meanings using the Nelson et al. (1999) word association norms. For instance, for the critical homograph ‘‘fall,’’ the 12-item set contained 6 words related to the ‘‘stumble’’ meaning and 6 words

Table 1
Experimental conditions used in Experiments 1–6

	6-Item (Exp. 1–6)	12-Item (Exp. 1)	12-Item (Exp. 2–6)
Homograph			
	STUMBLE	STUMBLE	STUMBLE
	SLIP	SLIP	**AUTUMN**
"fall"	RISE	RISE	SLIP
	TRIP	TRIP	**SEASON**
	FAINT	FAINT	RISE
	CLUMSY	CLUMSY	**SPRING**
	or	*"fall"* **AUTUMN**	*"fall"* TRIP
		SEASON	**LEAVES**
	AUTUMN	**SPRING**	FAINT
	SEASON	**LEAVES**	**BRISK**
	SPRING	**BRISK**	CLUMSY
"fall"	**LEAVES**	**HARVEST**	**HARVEST**
	BRISK		
	HARVEST		
DRM			
	SNOOZE	SNOOZE	SNOOZE
	WAKE	WAKE	**SLUMBER**
"sleep"	BEDROOM	BEDROOM	WAKE
	UNCONSCIOUS	UNCONSCIOUS	**LAY**
	DEEP	DEEP	BEDROOM
	BLANKET	BLANKET	**MOTEL**
		"sleep" **SLUMBER**	*"sleep"* UNCONSCIOUS
	or	**LAY**	**TRANCE**
		MOTEL	DEEP
	SLUMBER	**TRANCE**	**LAZY**
	LAY	**LAZY**	BLANKET
	MOTEL	**NIGHTMARE**	**NIGHTMARE**
"sleep"	**TRANCE**		
	LAZY		
	NIGHTMARE		

related to the "autumn" meaning. All 18 homograph sets are presented in the Appendix. Overall, the average BAS from each word to the critical item was .129, meaning approximately 13% of respondents in the Nelson et al. norms would respond with the homograph word when given a particular item and told to respond with "the first word that comes to mind." The total BAS from the set items to the critical homograph was equated across the different 6-item sets (.130 vs. .127, $p > .90$).

In some cases, we were only able to find 4 or 5 items related to a particular meaning of a homograph listed in the Nelson et al. norms. When this occurred, an additional word that was judged to be semantically related (e.g., harvest) was selected. These 16 items (out of 216 items total) are designated by an asterisk (*) in the Appendix. In a later norming procedure with 25 Washington University undergraduates, we presented these additional items on a sheet of paper with the instructions "please respond to each item by writing down the first word that comes to mind." Our results indicated that these additional items contained a BAS of around .18, a value not significantly different from the other 200 items obtained from the Nelson et al. (1999) norms ($p > .25$).

Each participant received 12 6-item sets and six 12-item sets. For the 6-item sets, all words were related to either the first meaning (e.g., "stumble") or the second meaning (e.g., "autumn") of the homograph (see Table 1, Column 1). For the 12-item sets, both meanings were combined, but presented in a grouped fashion for Experiment 1 (see Table 1, Column 2). For the 6-item sets, half of the participants received words related to the first meaning and half received words related to the second meaning. When these 6-item sets were combined for the 12-item condition, half of the participants received meaning 1 followed by meaning 2 whereas the other half received the opposite order. Finally, the sets were counterbalanced across participants such that each set appeared in the 12-item condition for one-third of the participants.

DRM item-set construction. For the DRM sets, we selected 18 critical items from the Stadler et al. (1999) norms. For each critical item, we then selected 12 of the Stadler et al. associates with the constraint that the average BAS of each item to the critical word would approximate the BAS of the homograph items. Because of the difficulty in finding enough high associates in the homograph sets, several of the highest associates in the Stadler et al. DRM lists were not selected. The resulting item-sets had an average BAS of .103. This value did not differ from the .129 BAS for the homograph item-sets [$t(34) = 1.39$, $SE = .02$, ($p > .15$).[1]

Each 12-item DRM set was broken into two 6-item sets. Because all of the words in these DRM sets were related to the same meaning, the assignment of each word to either of the two 6-item sets was arbitrary, with the constraint that the two halves must be equal in BAS. As with the homograph lists, each participant saw 12 6-item sets and six 12-item sets. For the 6-item sets, half of the participants received the first set (e.g., the "snooze" set, see Table 1) and half received the second set (e.g., the "slumber" set). In addition, for the 12-item set, half of the participants received set 1 followed by set 2 whereas the other half received the opposite order. Finally, the 12 6-item sets and the six 12-item sets were counterbalanced across participants such that each list appeared in the 12-item condition for one-third of the participants.

Study list structure. For each study trial, participants received 12 homograph related items, 12 DRM-related items, and 4 unrelated buffer items (2 at the beginning and 2 at the end of the list, to control for primacy and recency effects, respectively). The 12-DRM and 12-homograph items could either come from the same 12-item set or instead from two different 6-item sets. All possible combinations were included so that either a 12-item DRM list or two 6-item DRM lists could be paired with either a 12-item homograph list or two 6-item homograph lists. These combinations were counterbalanced across participants. In addition, the order of either the DRM or the homograph items within a study list was alternated across lists such that DRM items and homograph items appeared equally often in both the beginning and end of the study list.

Procedure

Participants were tested either individually or in pairs. They were seated approximately 60 cm away from a VGA monitor, read a set of task instructions displayed on the monitor and then heard them paraphrased by the experimenter. The participants were instructed that they would see 28 words shown for approximately 1.5 s each and to try to remember the words for a later recall test. Stimuli were always centered on the display monitor and presented in uppercase letters to reduce any effect of orthographic distinctiveness. Each stimulus word was preceded by a 500 ms inter-stimulus interval and presented for 1200 ms. After all 28 words were presented, participants wrote down as many words as they could remember on a separate recall sheet. After 2 minutes of attempting to recall the items, they were asked to prepare for the next series of words. This procedure continued for 12 total lists. Participants were allotted rest breaks as needed between each recall attempt and the following study list.

Design

There were three within-subject variables in the design: Item Type (DRM vs. homograph), Number of Related Words (6 vs. 12), and Study Status (studied item vs. nonstudied critical item). List items were taken from either DRM or homograph sets that contained either 6 or 12-related words. Participants' free recall was examined after each list.

Results

Veridical recall

A 2 (Item Type) × 2 (Number Related) ANOVA was used to compare recall for DRM and homograph list items presented in either the 6-related or 12-related condition. These percentages are shown in the left panel of Fig. 3. In general, veridical recall of items from the homograph list (e.g., autumn) was better than recall of the DRM items (e.g., slumber)—$F(1, 35) = 68.31$, $MSE = 36.4$. There was no effect of the number of related items. However, as predicted, there was a significant interaction between Item Type and Number of Related Items—$F(1, 35) = 8.13$, $MSE = 39.4$. The DRM items were recalled 4.7% ± 2.9% ("±" equals the 95% confidence interval) more accurately in the 12-related condition compared to the 6-related condition, whereas homograph items showed no such benefit. Hence, these data are consistent with the idea that subjects organized the items by meaning, and that such organization was beneficial to veridical recall primarily for the DRM items.

False recall

A 2 (Item Type) × 2 (Number Related) ANOVA was used to compare recall for the DRM and homograph

[1] When we later included the 16 items not included in the Nelson et al. (1999) norms, the average association strength of 13.5% from the homograph lists was marginally greater then the 10.3% value from the DRM lists [$t(34) = 1.76$, $SE = .02$, $p > .05$]. However, this should not be a problem because in the current series of studies we are primarily interested in patterns of *increase* in false memory as the number of associates increases from 6 to 12, rather than in making absolute comparisons between memory for homograph and DRM items per se.

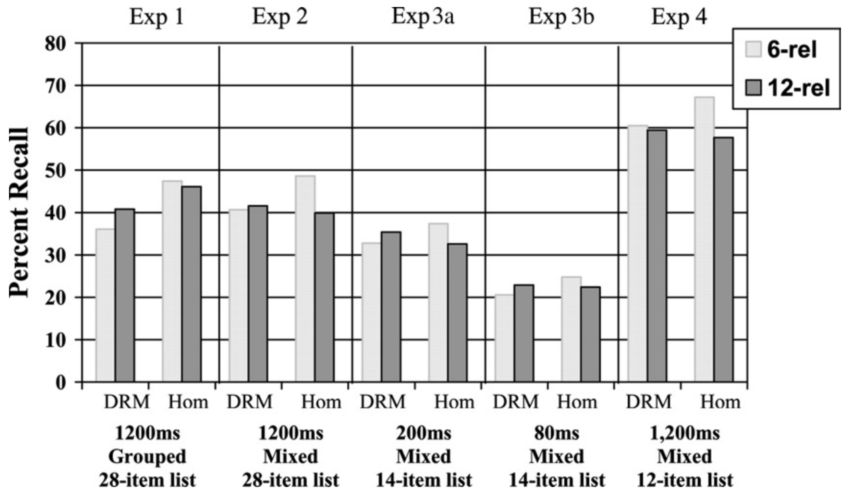

Fig. 3. Veridical recall of list items from the unambiguous (DRM) lists and the ambiguous (Hom) lists in Experiments 1, 2, 3a, 3b, and 4.

critical nonpresented items following either the 6-related or 12-related study lists. These percentages are shown in the left panel of Fig. 4. First, as noted there is very little difference in the false recall across the DRM and Homograph six-item lists. Hence, we were successful in equating false recall in our baseline sets. Contrary to veridical recall, there was no effect of Item Type on false recall of critical items. Clearly, the number of related words produced a strong effect on false recall of critical items. Adding 6-related words significantly increased false recall from 5.6% following 6-related words to 15.5% following 12-related words. More importantly, this significant $9.9\% \pm 4.3\%$ increase in false memory was remarkably similar for the homograph items $(10.4\% \pm 4.7\%)$ and the DRM items $(9.5\% \pm 5.8\%)$. These observations were supported by a significant main effect of the number of related words [$F(1,35) = 22.83$, $MSE = 155.5$], but no effects of either the type of item or the interaction of item type and number of related words (both F's < 1).

Discussion

Overall, these data are most consistent with the activation theory of false memory. Specifically, we obtained a functional dissociation of the effect of thematic coherence on veridical and false recall. There was an increase in veridical recall for DRM items when presented in the 12-related condition yet no corresponding increase for homograph items. However, such effects of meaningful

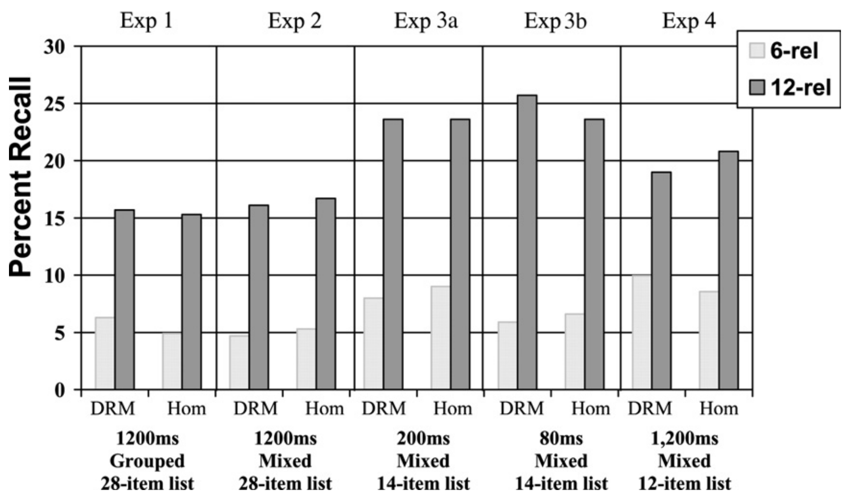

Fig. 4. False recall of unambiguous (DRM) and ambiguous (Hom) critical items in Experiments 1, 2, 3a, 3b, and 4.

organization did not influence false recall. Instead, false recall appeared to be driven by associative activation. We received the same increase in false memory regardless of whether the 6 additional words added to a list were of the same meaning or a different meaning.

These results are consistent with an activation account of false memory (Deese, 1959; Roediger, Balota et al., 2001; Underwood, 1965) in which the probability that a person accepts a critical nonpresented item as studied is primarily influenced by the extent to which it was implicitly activated by the studied list items. As mentioned previously, the current dissociation of meaning on veridical and false memory is inconsistent with thematic accounts of false memory, in which effects of thematic organization should be at least as great in false memory as in veridical memory.

Experiment 2

One potential concern about Experiment 1 is the possibility that grouping the homograph sets created a conscious identification of the "linking" word. For instance, when studying the items "*stumble*, *slip*, *rise*, *trip*, *faint*, *clumsy*, *autumn*, *season*, *spring*, *leaves*, *brisk*, *harvest*", the word "autumn" might cue people to consciously think of "fall" as the word that links the two sets of items. Although such a process should generally reduce false memories to "fall" if it is realized that the word itself was never presented, false memories may be produced if people forget the source of the word's familiarity at test. In order to reduce this problem in the second experiment (and remaining experiments), we alternated the two meanings in the 12-related condition (see Table 1, Column 3). The 6-item conditions were identical to Experiment 1.

Alternating list items also allowed us to make a stronger test of the effect of thematic coherence on both veridical and false memory. In particular, it is possible that our observed rate of false memory in the 12-item Homograph lists were driven by 2 independent 6-item themes that were each 1/2 as strong as the single theme derived from the 12-item DRM lists, thus producing the same increase in false memory when combined. In contrast, for list items, the two separate themes present in the 12-item homograph list would not benefit recall relative to the two separate themes in the 6-item condition. Experiment 2 was designed to also provide a test of this alternative account. Specifically, alternating list items in the 12-item condition should decrease veridical recall for the homograph items relative to the blocked 6-item condition, by disrupting people's ability to construct two organized themes. Therefore, a blocking effect is predicted in which alternation of meanings should hurt veridical recall of homograph lists in Experiment 2, relative to both the DRM list items in Experiment 2 and the blocked presentation of homograph lists in Experiment 1. This should occur because alternating meanings should hurt participants' ability to semantically organize the homograph list items into 2 coherent themes during study. However, if false recall is due primarily to BAS, then we would predict no difference in false recall between the critical homograph items in Experiments 1 and 2. Similarly, we are predicting no difference in false recall between the critical homograph items and critical DRM items in Experiment 2. Comparing homograph list performance in Experiment 2 to Experiment 1 therefore allows us to compare the effects of blocking on both veridical and false recall while holding all the items constant, the only difference being the manner in which the list items were presented: blocked vs. alternated. If thematic coherence drives false recall, then a similar effect of blocked vs. alternated presentation should be observed in veridical and false recall. Alternatively, if false recall is driven primarily by lexical associative activation, homographic critical items should continue to show the same increase in false recall as both DRM critical items in Experiment 2 and homograph critical items in Experiment 1.

Method

Participants

Thirty undergraduates at Washington University in Saint Louis participated for partial completion of a research requirement for an introductory psychology class. All were native English speakers with normal or corrected-to-normal vision.

Stimuli

The stimuli in Experiment 2 were identical to those used in Experiment 1 with one exception. The items in the 12-related conditions alternated between the two separate item-sets. As shown in Table 1, for the homograph sets, this involved switching meanings between every item such that the set "*stumble*, *slip*, *rise*, *trip*, *faint*, *clumsy* " and the set "*autumn*, *season*, *spring*, *leaves*, *brisk*, *harvest*" would combine to form the 12-item set "*stumble*, *autumn*, *slip*, *season*..." For the DRM items, the two sets were also alternated such that the set "*slumber*, *lay*, *motel*, *trance*, *lazy*, *nightmare*" and the set "*snooze*, *wake*, *bedroom*, *unconscious*, *deep*, *blanket*" would combine to form the 12-item set "*slumber*, *snooze*, *lay*, *wake*..."

Design and procedure

The design and procedure of Experiment 2 were identical to Experiment 1.

Results

Veridical recall

As shown in Fig. 3, the veridical recall performance in Experiment 2 reveals a different pattern from that of the first experiment. As predicted, intermixing the two sets in the 12-related condition impaired recall for the homograph sets, but did not affect performance on the DRM sets. A 2 (List Type) × 2 (Number Related) ANOVA confirmed this conclusion, revealing a significant interaction of List Type × Number Related — $F(1,29) = 10.83$, $MSE = 64.2$. Recall of homograph items was impaired by 8.7% ± 4.3% in the alternated 12-related condition, recall of DRM items showed a nonsignificant increase (1.0% ± 4.3%) between the grouped 6-related and the alternated 12-related presentation conditions.

False recall

In contrast to the veridical recall data, the false recall data mirrored that of Experiment 1. Specifically, as shown in Fig. 4, there is little difference in the false recall across the DRM and Homograph six-item lists. Adding 6-related words increased false recall from 5.0% following 6-related words to 16.4% following 12-related words. As with Experiment 1, this significant 11.4% ± 4.0% increase in false recall was virtually identical for the homograph critical items (11.4% ± 5.8%) and the DRM critical items (11.4% ± 6.4%). These observations were substantiated by a significant main effect of the number of related words [$F(1,29) = 33.19$, $MSE = 117.3$], but no effects of either the type of item or the interaction of item type and number of related words (both Fs < 1).

Experiments 1 and 2 combined

A 2 (Number Related) × 2 (Experiment) mixed ANOVA was used to address whether veridical memory for the homograph list items was indeed impaired by intermixing the meanings within the study list. This analysis revealed a significant interaction between Experiment and number of related items [$F(1,64) = 7.95$, $MSE = 56.5$], showing impaired memory for 12 intermixed items compared to the 6 blocked item list (Experiment 2), but no such impairment when the 12-item lists were grouped separately by meaning (Experiment 1). This same analysis performed on false memories revealed an effect of the number of related items [$F(1,64) = 35.87$, $MSE = 108.0$], but no hint of an interaction with experiment ($F < 1$). Thus, even within homograph lists, a manipulation designed to decrease thematic coherence showed a dissociative effect on veridical and false recall.

Discussion

These data provide compelling evidence that the thematic coherence of a list is not essential in producing false memory. In Experiment 2, we directly manipulated the organization of meanings within the study list. As predicted, disrupting thematic coherence during study impaired participants' veridical memory. In particular, participants recalled fewer studied items from the homograph list when the items alternated across meanings than when the meanings were grouped. Importantly, even though such organization influenced veridical memory, it had no effect on false memory. As in Experiment 1, participants showed as large of an increase in false memory when the additional 6 studied words were related to a different meaning as when they referred to the same meaning. Thus, we have evidence that the 2 themes derived from the 12-item intermixed homograph lists in Experiment 2 are NOT as strong as each 6-item theme. Nonetheless, we still received an equal increase in false memory.

The data from both Experiments 1 and 2 are consistent with an associative activation account in which false memories arise due to the associative activation of the critical item by the list items during study. The strong prediction from this associative theory, that both DRM and homograph critical items would show the same increase in false recall when the two 6-item subsets were combined, was confirmed in Experiments 1 and 2. In contrast, these data are inconsistent with a theme-based account of false memory in which false memories are driven by the underlying theme or gist of the studied items. Such a theory cannot explain the dissociation we observed regarding effects of same vs. different meanings on veridical and false recall. Nor could such a theory account for our blocking effect for homograph lists, with intermixed lists disrupting only list items (presumably through disrupting thematic organization) but having no effect on false recall of critical lures. However, before accepting this conclusion, we decided to reduce the presentation duration to see if the results would generalize to situations with little time for the encoding of item-specific information.

Experiment 3a and 3b

In addition to the lexical activation or meaning extraction of an item, researchers have proposed that participants encode item-specific information during study (see Balota et al., 1999; McDermott & Watson, 2001; for discussion). For example, one may remember the appearance of the item or specific contextual details surrounding the presentation of the item. According to Balota et al. (1999); accurate memory in the DRM paradigm requires the correct discrimination at test between item-specific information activated during study and highly activated lexical/semantic representations of nonpresented critical items. At relatively long presentation durations, item-specific information should be helpful in discriminating veridical from false memories, since veridical memories would contain stronger item-specific details than false memories. At shorter durations, however, less item-specific information would be available forcing participants to rely more heavily on overall lexical activation or meaning. In order to minimize effects of item-specific processing in Experiment 3, we presented the study items for either 200 or 80 ms. In addition, this decreased duration should further reduce the concern from Experiment 1 in which critical "linking" words for the homograph lists may consciously come to mind during study.

Method

Participants

Forty-eight undergraduates at Washington University in Saint Louis participated for partial completion of a research requirement for an introductory psychology class. All were native English speakers with normal or corrected-to-normal vision. Twenty-four of the participants received the 200 ms presentation duration (Experiment 3a) and 24 received the 80 ms presentation duration (Experiment 3b).

Stimuli

The study lists in Experiment 3 were identical to those used in Experiment 2 with one exception: in Experiments 1 and 2 participants studied 12 lists composed of 28 words each, however, in Experiment 3 these lists were divided such that each participant studied 24 lists of 14 words each. This was done to reduce possible floor effects in veridical recall. The same buffer words were used as in the first two experiments, however, only one word appeared at the beginning and end of each list.

Design and procedure

The design and procedure of Experiment 3 were identical to the previous experiments with the following exceptions: first, the participants now completed 24 recall attempts (of 14 words) instead of 12 (of 28 words). Second, the items were presented at a much faster rate than in the previous experiments. In Experiment 3a, the items were presented for 200 ms each with 50 ms ISIs and in Experiment 3b the items were presented for 80 ms each with 32 ms ISIs. Therefore, in Experiment 3a the entire list was presented in 3.5 s and in Experiment 3b the entire list was presented in 1.6 s. Participants were instructed that the items would be flashed very quickly and asked to remember as many of the words as they could.

Results

Veridical recall

In both Experiments 3a and 3b, the veridical recall performance mimicked that shown for the 1200 ms stimuli. In Experiment 3a, there was a $4.8\% \pm 2.8\%$ impairment in recall for homograph items in the intermixed 12-related condition relative to the 6-related condition, yet DRM items showed an increase of $2.7\% \pm 2.2\%$. Similarly, in Experiment 3b, there was a $2.4\% \pm 1.6\%$ impairment in recall for homograph items, yet an increase of $2.3\% \pm 2.0\%$ for the DRM items. These effects were supported by significant Item Type × Number Related interactions in both experiments—$F(1,23) = 32.18$, $MSE = 10.4$; $F(1,23) = 27.32$, $MSE = 4.9$, for Experiments 3a and 3b, respectively.

When the data from participants in the 2 presentation conditions were combined, the Type × Number Related interaction was highly significant—$F(1,46) = 54.00$, $MSE = 7.8$. There was also a main effect of presentation duration [$F(1,46) = 64.08$, $MSE = 105.7$] with greater recall for list items presented for 200 ms than 80 ms. However, presentation duration did not participate in any higher order interactions (all p's $> .10$).

False recall

The false recall data from Experiments 3a and 3b also mirrored those of Experiments 1 and 2, even though participants were given very little time for item-specific encoding or conscious linking of the nonpresented critical item. Specifically, as shown in Fig. 4, there is very little difference in the false recall across the DRM and Homograph six-item lists. In addition, adding 6-related words significantly increased false recall from 8.5% (Experiment 3a) and 6.3% (Experiment 3b) following 6-related words to 23.6% (Experiment 3a) and 24.7% (Experiment 3b) following 12-related words. Under 200 ms presentation conditions, the homograph critical items showed a $14.6\% \pm 6.5\%$ increase in false memory and the DRM critical items showed a $15.6\% \pm 9.0\%$ increase in false memory with 6 additional related words. Similarly, under the 80 ms presentation conditions, the increases were $17.0\% \pm 8.3\%$ for homograph items and $19.8\% \pm 8.5\%$ for DRM items.

When the false recall data from participants in the 2 presentation duration conditions were combined, the only significant effect was a main effect of the number of related items—$F(1,46) = 48.54$, $MSE = 277.6$. DRM and homograph items showed identical patterns of false recall, as indicated by the lack of a main effect of Item Type and no interaction between Item Type and Number Related (both F's < 1).

Discussion

For the studied items, there was still an effect of alternating meanings for the homograph condition even at a presentation duration of 80 ms. However, as with the previous studies, adding 6 associated items during study increased false recall equally for both types of items, regardless of whether the extra 6 items converged on the same or different meaning. This pattern of equal false memory combined with differences in veridical memory occurred at presentation rates that should have reduced item-specific encoding and conscious identification of possible critical items. This pattern of results further supports an associative activation theory of false memory.

Experiment 4

Unfortunately, Experiments 3a and 3b differed from the earlier experiments in ways other than the presentation duration. Specifically, to prevent floor effects due to the rapid presentation of items in Experiment 3, 24 14-item lists were presented to participants rather than the 12 28-item lists presented in Experiments 1 and 2. This meant that the DRM and Homograph lists in Experiment 3 were presented separately, rather than combined in each study trial. In order to show that this procedural differences did not influence the pattern of results, we decided to use the 24 12-item lists from Experiments 3a and 3b and present those lists for the 1200 ms per-item duration used in Experiments 1 and 2.

The fourth experiment also addressed another methodological concern in Experiments 1–3. As mentioned previously, these lists began and ended with either 1 (Experiments 3a and 3b) or 2 (Experiments 1 and 2) buffer words included to minimize primacy and recency effects in free recall. However, the use of buffer words probably reduced the overall thematic coherence of the lists and may have discouraged semantic organization by participants. (Note, however, that this criticism cannot account for the blocking effects obtained in Experiments 1 and 2.) In order to show the generality of our results to those in most standard false memory experiments, we decided to replicate the intermixed presentation results of Experiment 2 while eliminating the buffer trials.

Method

Participants

Thirty-six undergraduates at Montana State University participated for partial completion of a research requirement for an introductory psychology class. All were native English speakers with normal or corrected-to-normal vision.

Stimuli

The study lists in Experiment 4 were identical to those used in Experiment 3 with one exception: the buffer items were removed from the beginning and end of the lists.

Design and procedure. The design and procedure of Experiment 4 were identical to Experiment 3 except that the items were presented for 1200 ms each.

Results

Veridical recall

As shown in Fig. 3, the pattern of veridical recall performance in Experiment 4 is virtually identical to Experiment 2 despite the higher overall recall rate in Experiment 4 (61%) relative to Experiment 2 (42%). As was found for Experiment 2, intermixing the two sets in the 12-related condition impaired recall for the homograph sets, but did not affect performance on the DRM sets. A 2 (List Type) $\times$ 2 (Number Related) ANOVA confirmed this conclusion, revealing a significant interaction of List Type $\times$ Number Related—$F(1,29) = 30.81$, $MSE = 20.6$. Recall of homograph items was impaired by $9.5\% \pm 2.2\%$ in the alternated 12-related condition, recall of DRM items showed a nonsignificant decrease ($1.1\% \pm 2.8\%$) between the grouped 6-related and the alternated 12-related presentation conditions.

False recall

The pattern of false recall also mirrored that of Experiment 2, despite showing higher overall levels of false recall (14.6 and 10.1% for Experiments 4 and 2, respectively). Specifically, there was again little difference in the false recall across the DRM and Homograph six-item lists and the increase in false recall was similar for the homograph critical items ($12.3\% \pm 6.3\%$) and the DRM critical items ($9.0\% \pm 6.7\%$). These observations were substantiated by a significant main effect of the number of related words [$F(1,35) = 18.66$, $MSE = 220.2$], but no effect of the type of item nor the interaction of item type and number of related words (both F's < 1).

Discussion

The results of Experiment 4 replicated the pattern of data in Experiment 2 despite (1) eliminating the buffer items and (2) presenting 12 items per list rather than 28. Therefore, our recurrent dissociation between effects of thematic coherence on veridical and false memory is not simply an artifact of our use of buffer items or combined lists in Experiments 1 and 2. The final two experiments will further expand upon the generality of our results by using a recognition, rather than recall task.

Experiment 5

In the previous 4 recall experiments, the data suggest a very limited role for thematic coherence in creating false memory. However, that does not eliminate the possibility that meaning would play a role during a recognition task. When presented with an item and asked whether it was studied or not, people could rely on the semantic relatedness to studied items when making their decision. Indeed, some researchers have suggested that there would be an increased use of semantic similarity during a recognition test (Gallo & Roediger, 2002; Miller & Wolford, 1999), compared to a recall test. In addition to the ''old/new'' judgment, the use of a recognition test also allows for the use of the remember–know procedure developed by Tulving (1985) to examine the possibility that Homograph and DRM items elicit subjectively different basis for responding. Previous researchers have provided dissociations between ''remember'' and ''know'' responses (see Gardiner, Ramponi, & Richardson-Klavehn, 2002; for a recent review), supporting the argument that they reflect qualitatively distinct phenomenological states. To the extent that false memories differ between the two types of lists, one might expect to find different patterns of ''remember'' vs. ''know'' responses for DRM and Homograph critical items. In particular, we predict that alternating the meanings of homograph list items should reduce the amount of relational processing during study (indeed, we have been making this argument all along). Interestingly, several researchers have made reference to relational and/or distinctive processing in false memory (Arndt & Reder, 2003; Dodson & Schacter, 2002; Schacter, Cendan, Dodson, & Clifford, 2001). For instance, Arndt and Reder (2003) manipulated the similarity in fonts of list items corresponding to each theme during study. Arndt and Reder found that presenting each item within a theme in its' own unique font (argued to reduce relational processing) reduced false memory, yet had little effect on veridical memory (i.e., ''hit'' rates). Further, this reduction was primarily reflected in ''remember'' responses. This pattern occurred regardless of whether the unique vs. common font variable was manipulated within or between participants, suggesting the effect was not due to participants' simply raising their criterion for an ''old'' response by requiring more detailed recollection (see Schacter, Israel, and Racine for a discussion of such a ''distinctiveness heuristic''). Arndt and Reder argued that relational processing enhances the activation of ''theme nodes'' during study, producing greater false memory for thematically consistent lures. However, this pattern could also be explained by the critical lure reminding participants of a studied item presented in that same font. This would allow the participants to use what has been called a ''recall-to-reject'' strategy (see Rotello, Macmillan, & Van Tassel, 2000; for a discussion) because recalling the studied item in the same font would allow them to reject the critical lure. This strategy would only work in the unique font condition because recalling a related item in the common font condition does not allow for the rejection of the lure. This could certainly explain a reduction in ''remember'' responses in the unique font condition.

Although his focus was on veridical memory, a study by Mantyla (1997) is useful here because he directly manipulated relational vs. distinctive encoding instructions while holding the presentation stimuli constant. Participants in Mantyla's experiment were told to focus on either the similarities (relational processing) or differences (distinctive processing) among studied faces. Of interest, Mantyla observed a dissociation between type of processing and memory judgments. Specifically, ''know'' responses were higher following relational processing than following distinctive processing while ''remember'' responses were higher following distinctive processing.

Based upon the Mantyla (1997) study, we predict that hit rates should be differentially affected by our intermixed 12-item list for DRM and Homograph lists. For DRM lists, both the 6-item blocked and 12-item intermixed conditions should encourage relational (rather than item-specific) processing because the items continue to converge upon the same theme. However, for the Homograph lists, we would expect to observe a reduction in ''hit'' rates in the intermixed 12-item condition relative to the blocked 6-item condition mimicking that found in recall (corresponding to a drop from relational to item-specific processing). Moreover, if indeed this drop is due to less relational processing, than this drop should occur in the ''know,'' rather than ''remember'' judgments. For false memories, the thematic and activation accounts make different predictions. If false memories are primarily driven by relational processing, than we might expect a pattern similar to that found by Arndt and Reder (2003) in which reducing relational processing should reduce false memory and this should be driven primarily by ''remember'' judgments (assuming their results were indeed due to less relational processing rather than a ''recall-to-reject'' strategy). If, however, false memories are driven by implicit associative activation, than we should continue to find the same increase in false memory for Homograph critical lures as for DRM critical lures in the 12-item intermixed condition.

Method

Participants

Sixty-four male and female undergraduates at Montana State University participated for partial completion of a research requirement in an introductory psychology class. All were native English speakers with normal or corrected-to-normal vision.

Stimuli

The stimulus lists in Experiment 5 were identical to those used in Experiment 2. Following the presentation of each 28-word list, participants received a recognition test consisting of 8, 12, or 16 items. It was necessary to vary the length of the recognition test in this way because the studied lists (though equated in length) contained 2, 3, or 4 themes. We therefore needed to vary the length of the recognition test to present an equal number of items per theme. Each recognition test contained 1 list item, 1 critical item, 1 low associate, and 1 unrelated item for each 6 or 12-item list. If the study list contained four 6-item lists (2 Homograph and 2 DRM), than the recognition test would consist of 16 words (4 studied and 12 lures). In contrast, if the study list contained two 12-item lists (1 Homograph and 1 DRM) than the recognition test would consist of 8 words (2 studied and 6 lures). For the other two possible combinations in which there is one 12-item list and two 6-item lists, the recognition test would consist of 12 words (3 studied and 9 lures). The studied associates were taken from the fourth position of the original sublist. For example, given the sublists "*stumble*, *slip*, *rise*, *trip*, *faint*, *clumsy*" and "*slumber*, *lay*, *motel*, *trance*, *lazy*, *nightmare*," the selected associates in the 6-item conditions would be "trip" and "trance." In the alternated 12-item condition, these same fourth associates were selected. For instance, given the combined list "*stumble*, *autumn*, *slip*, *season*, *rise*, *spring*, *trip*, *leaves*, *faint*, *brisk*, *clumsy*, *harvest*," the selected associate would be "trip" and given the combined list "*autumn*, *stumble*, *season*, *slip*, *spring*, *rise*, *leaves*, *trip*, *brisk*, *faint*, *harvest*, *clumsy*," the selected associate would be "leaves." The low associate items were words that were associated to at least one word in the study list, but were only weakly associated or not associated to the critical item. Care was taken to ensure that the low associates for the Homograph and DRM lists did not differ either in their associations to a specific list item (.10 vs. .09, for DRM and Homograph items, respectively) or to their association to the critical nonpresented item (.02 vs. .01, for DRM and Homograph items, respectively). Finally, 36-unrelated words were chosen that did not contain strong associations to any studied words. The items on the recognition test were presented randomly.

Design and procedure

For both the Homograph lists and the DRM lists, there were four types of words presented during the recognition tests: critical nonpresented words, studied fourth associates, nonstudied low associates, and nonstudied unrelated words. As in Experiments 1, 2, and 4 each item was presented for 1200 ms during study with a 500 ms ISI. During each recognition test, participants were instructed to respond quickly, but not at a cost to accuracy, by pressing either the "p" key (labeled as "o" for "old") or the "q" key (labeled as "n" for "new"). They were instructed that following a "new" response, they would be asked to press the space bar to continue. They were also instructed that, following an "old" response, they would be further prompted to make a remember–know–guess distinction. They were asked to press the "h" key (labeled as "r" for "remember") if "you have a conscious recollection of the word from the study list, such as the details of its appearance or a recollection of what you were thinking about at the time." They were asked to press the "g" key (labeled as "k" for "know") if "you are sure the item was presented, but cannot recollect its actual occurrence or any related details." Finally, they were told to press the "f" key (labeled as "g" for "guess") if "you neither remember nor know the item was presented, and are merely guessing that it was presented."

Results

The percentage of "old," "remember," "know," and "guess" responses to studied items, critical nonpresented items, low associate items, and unrelated items are shown in Table 2.

Veridical memory (hits)

There were no significant effects of either list type or number of related associates to "old" responses (all Fs < 1). However, as predicted, there was a significant $7.0\% \pm 5.6\%$ drop in "know" responses for Homograph list items in the 12-item intermixed list [$t(63) = 2.51$], with no corresponding drop for DRM items ($0.9\% \pm 5.7\%$), though the Item Type $\times$ Number Related interaction in "know" responses failed to reach significance ($p < .11$). No other effects approached significance.

False memory (false alarms)

As was so for the recall experiments, there was a large effect of the number of related items on critical item false alarms, with participants making $10.6\% \pm 3.8\%$ more false alarms following the study of 12-related items than following the study of 6-related items—$F(1,63) = 30.78$, $MSE = 234.5$. In contrast with recall, there was a main effect of Item Type [$F(1,63) = 10.41$, $MSE = 237.2$], with $6.2\% \pm 3.9\%$ more false alarms to Homograph than to DRM critical items. However, the $11.2\% \pm 6.1\%$ increase in false alarms for DRM items in the 12-related condition did not differ from the $10.0\% \pm 5.0\%$ increase for Homographs ($F < 1$). This data pattern remained the same in both the "remember" and "know" judgments (both Fs < 1 for the Item Type $\times$ Number Related interaction). When the "remember" and "know" judgments were combined, the $10.2\% \pm 5.4\%$ increase in false alarms for DRM items did not differ from the $10.2\% \pm 4.9\%$ increase for Homographs ($F < 1$). In contrast to the critical items, the low associate items did show an Item Type $\times$ Number Related interaction—$F(1,63) = 5.21$, $MSE = 201.7$. Specifically, the $7.6\% \pm 4.4\%$ increase in false alarms for DRM items in the 12-related condition differed from the null effect for Homographs ($0.5\% \pm 4.9\%$). This is as would be expected, since only 6 of the 12-items for

Table 2
Recognition results in Experiment 5

	DRM				Homograph			
	"old"	R	K	G	"old"	R	K	G
Studied								
6-related	.78	.26	.39	.12	.79	.28	.39	.11
12-related	.78	.29	.38	.11	.78	.31	.32	.15
Difference	*.00*	*.03*	*−.01*	*−.01*	*−.01*	*.03*	*−.07**	*.04*
Critical items								
6-related	.33	.10	.09	.14	.40	.12	.11	.16
12-related	.44	.15	.14	.14	.50	.20	.14	.16
Difference	*.11***	*.05**	*.05***	*.00*	*.10***	*.08***	*.03*	*.00*
Low associates								
6-related	.12	.04	.02	.06	.18	.05	.03	.10
12-related	.19	.05	.03	.11	.17	.06	.03	.08
Difference	*.07***	*.01*	*.01*	*.05**	*−.01*	*.01*	*.00*	*−.02*
Unrelated			"old"	R	K	G		
			.06	.02	.00	.04		

* $p < .05$.
** $p < .01$.

homograph lists would be associatively and/or semantically related to the low associate. Nonetheless, this interaction is important in that it demonstrates that the lack of an interaction for our critical items was not due to a lack of statistical power. Finally, there was only a 6% false alarm rate to unrelated words. The fact that the hit rate for list items was lower than the correct rejection rate for unrelated items indicates a general bias to respond "new." This bias is appropriate since 75% of the recognition items were lures.

Discussion

The results of Experiment 5 extended the findings from recall into the realm of recognition. The main pattern of results remained. Specifically, homograph and DRM critical items showed the same increase in false memory in the intermixed 12-item list, despite the fact that the extra 6 studied associates were of a completely different meaning in the homograph lists. This "null" Item Type × Number Related interaction persisted even in the face of a significant interaction for our low associate items, thus demonstrating that differences in thematic organization did exist between the intermixed DRM and Homograph lists. The recognition data allowed us to examine possible phenomenological differences in responding between the two groups of items. Interestingly, intermixing the meanings in the homograph 12-item condition (thus forcing less relational processing) selectively decreased "know" responses while numerically (but not significantly) increasing "remember" responses. This finding is consistent with Mantyla's (1997) face encoding study discussed previously in which relational (relative to distinctive) processing selectively increased "know" responses. Alternating the meanings during study likely reduced the amount of relational processing between study items, rendering them less integrated with other members of the category. The fact that such processing is important for veridical memory is seen in the dropped recall and hit rates in this condition. However, this manipulation did not affect false recall or false alarms, suggesting that relational processing of list items is not necessary for the development of false memories.

Experiment 6

One possible criticism regarding each of the previous 5 experiments is that we have not yet conclusively demonstrated that the gist trace for 12-item DRM lists is actually stronger than the gist trace for 12-item Homograph lists. Instead, we have inferred this from the different patterns of performance for studied items between the DRM and Homograph lists. From this pattern, we have argued that adding 6 additional associates of the same meaning (i.e., DRM lists) creates a stronger gist trace than adding 6 associates of a different meaning (i.e., Homograph lists). In Experiment 6, we decided to directly test this assumption. Therefore, we replicated the procedure of Experiment 5 with one change: Participants were instructed to indicate how strongly related each item was to the other words in the list, rather than whether it was "old" or "new."

If indeed the 12-item DRM lists provide greater increases in gist than the 12-item Homograph lists, we would expect to see a greater increase in judged "relatedness" when 6 additional associates are studied for the DRM critical items than for the Homograph critical items.

Method

Participants

Thirty-eight undergraduates at Washington University participated for partial completion of a research requirement for an introductory psychology class. All were native English speakers with normal or corrected-to-normal vision. The data from 2 participants were eliminated for a failure to follow instructions.

Stimuli, design, and procedure

The study lists, design, and procedure in Experiment 6 were identical to those used in Experiment 5 with 1 exception: rather than judge whether each recognition test item was "old" or "new," participants were asked to judge "how closely related the test word is in meaning to the words actually studied." As with Experiment 5, participants were asked to respond quickly, but not at a cost to accuracy. They were asked to rate each word on a scale from 1 (not related at all) to 9 (very strongly related). Participants were asked to use the full range of the scale.

Results

The relatedness judgments to studied items, critical nonpresented items, low associate nonstudied items, and unrelated items are shown in Table 3.

Table 3
"Relatedness" decision from 1 to 9 and converted proportion scores[a] (in parentheses) for Experiment 6

	"Relatedness decision"	
	DRM	Homograph
Studied		
6-related	7.57 (.84)	7.66 (.85)
12-related	7.90 (.88)	7.68 (.85)
Difference	*+.33*[*]	*+.02*
Critical items		
6-related	6.98 (.77)	7.09 (.79)
12-related	7.93 (.88)	7.49 (.83)
Difference	*+.95*[**]	*+.40*[*]
Low associates		
6-related	4.96 (.55)	5.96 (.67)
12-related	6.35 (.70)	6.10 (.68)
Difference	*+1.39*[**]	*+.14*
Unrelated	2.74	

[a] Proportion scores were derived by dividing relatedness judgments by the maximum possible (i.e., "9").
[*] $p < .05$.
[**] $p < .01$.

Studied items

Similar to the earlier recall experiments, there was a marginally significant interaction between Item Type and number of related items—$F(1, 35) = 3.65$, $MSE = 22.5$, $p < .07$. This marginal interaction was caused by a significant .32 ± .31 increase in judged relatedness for DRM list items in the 12-item intermixed list coupled with no corresponding increase for Homograph items (.02 ± .29).

Nonstudied items

For the first time, the pattern of data for DRM critical items differed from that for Homograph critical items. Specifically, the increase in judged "relatedness" to list items was greater for DRM critical items (a .95 ± .32 increase) than for Homograph critical items (a .40 ± .36 increase). This numerical difference was substantiated by a significant interaction between Item Type and number of related items—$F(1, 35) = 4.39$, $MSE = 60.7$.

The low associate items also showed the critical Item Type × Number Related interaction—$F(1, 35) = 17.1$, $MSE = 81.9$. The 1.39 ± .54 increase in judged "relatedness" for DRM items in the 12-related condition differed from the null effect for Homographs (.14 ± .44). As with Experiment 5, this finding is not surprising, since only 6 of the 12-items for homograph lists would be associatively and/or semantically related to the low associate. Finally, unrelated items were given a rating of only 2.74 out of 9.

Discussion

The results of Experiment 6 are critical in demonstrating that our manipulations were effective in creating stronger "gist" traces for DRM critical items than for Homograph critical items. Of importance, even though the DRM critical items were judged as more consistent with the overall list theme, they showed no increase in false memory over less thematically consistent homographic critical items matched on associative strength. This pattern coupled with the results of the earlier experiments, suggests that thematic consistency plays a limited role at best in producing false memory in the DRM paradigm. When a critical nonpresented item comes to mind during a memory test, the fact that it is strongly related to the gist of the study list does not appear to be the driving force in accepting it as having been on the list. Instead, this acceptance appears more likely driven by associative activation processes that occur during encoding and/or retrieval.

General discussion

The present series of studies show that effects of meaning (at least as traditionally conceptualized) do not play an important role in false memory using the DRM paradigm. Across four recall experiments, the DRM and homograph critical items showed the same false recall both in the six item control condition and in the 12-item experimental condition. In fact, when the data from all 4 recall experiments were combined, the false recall for DRM and homograph critical items in the 6-related condition was 7.1 and 6.8%, respectively. Similarly, the false recall for DRM and homograph critical items in the 12-related condition was 19.4 and 19.5%, respectively. The 12.4% ± 3.1% increase in false recall for DRM items was virtually identical to the 12.8% ± 2.6% increase for homograph items. This pattern again emerged in the recognition experiment, with virtually identical increases in false alarms for DRM and Homograph critical items. Of importance, the "null" effects of List Type on false memory occurred even though the DRM critical items were shown to be more semantically similar than the Homograph critical items to their respective studied list items in Experiment 6. Thus, these data suggest that thematic coherence does not strongly modulate false memory in the traditional DRM paradigm.

This "null" effect of thematic coherence on false memory also occurred in conjunction with large effects of coherence on veridical memory. In every recall experiment there was a significant interaction between the number of associates and Item Type (this effect was also in the predicted direction for the recognition test after eliminating guesses, but failed to reach significance). With the grouped 12-item presentation of Experiment 1, this interaction reflected an increase in recall with the number of related items only when those extra items were of the same meaning (i.e., DRM items). Specifically, when twelve items of the same meaning were studied (i.e., DRM items), people recalled more words than when two different sets of six items were studied. For homograph items, when 6 items related to one meaning were combined with 6 items of another meaning, people recalled the same amount of words as when two different sets of six items were studied. This pattern would be predicted if people were unaware that the 2 sets of words were actually linked via a mediating homograph. In the intermixed experiments (Experiments 2–5), this interaction reflected either an increase (Experiments 3a and 3b) or no difference (Experiments 2 and 4) for DRM items, but a significant decrease in recall or recognition when two different meanings were alternated (i.e., homograph items). Moreover, when we examined the effect of blocking on homograph lists (blocked meanings in Experiment 1 vs. alternated meanings in Experiment 2), there was a significant blocking effect suggesting that intermixing the meanings disrupted participants' ability to semantically organize the list items during study. The recognition data in Experiment 5 further add to this explanation, suggesting that relational processing is particularly disrupted when meanings are intermixed rather than blocked. In summary, these data show robust effects of meaning on veridical recall and are consistent with the common assumption that meaningful organization and relational processing of study items improves subsequent memory performance.

Relation to Balota and Paul (1996)

The false memory data from the present experiment closely replicate the pattern of semantic priming discovered by Balota and Paul (1996). Across 5 experiments using either lexical decision or naming tasks, Balota and Paul found no difference in the pattern of priming between 2 primes that converged on the same meaning of an unambiguous target (e.g., lion-stripes-TIGER) or instead diverged unto 2 separate meanings of an ambiguous target (e.g., kidney-piano-ORGAN). It was not until they switched to a "relatedness judgment" task that Balota and Paul finally found a difference, with greater priming when the two primes converged upon the same meaning. Similarly, in our 5 experiments examining false memory using both recall and recognition tasks, we found no difference in false memory between our DRM lists that converged on the same meaning of an unambiguous critical item and our Homograph lists that diverged unto 2 separate meanings of an ambiguous critical item. As with Balota and Paul, it was not until we switched to a "relatedness judgment" task that we actually found a difference between the lists. These "null" effects on false memory combined with the significant difference in judged "relatedness" are shown in Table 4. The replication of Balota and Paul suggests that both semantic priming and DRM false memories are driven primarily by associative activation and not by degree of semantic relatedness. To obtain effects of relatedness, one may need to make semantic similarity part of the decision process itself.

Relation to other false memory studies

Robinson and Roediger (1997) found false recall percentages of 3, 19, and 25% using lists of 3, 6, and 12 associates, respectively (also see Watson et al., 2003). In contrast, we obtained 6.2% false recall in the 6-item condition and 19.4% false recall in our 12-item condition. It is likely that the smaller rates of false recall in the present study are due to the lower levels of backward associative strength in our lists relative to those used by Robinson and Roediger. It is interesting to note that the

Table 4
Increase in false memory in 12-item list (relative to 6-item list) for DRM and homograph critical items across Experiments 1–6

Experiment and target type	Experiments 1–6		
	n	Increase in FM (12 vs. 6) (%)	Difference
Experiment 1	36		
DRM		9.5	
Homograph		10.4	−0.9% ns
Experiment 2	30		
DRM		11.4	
Homograph		11.4	0.0% ns
Experiment 3a	24		
DRM		15.6	
Homograph		14.6	+1.0% ns
Experiment 3b	24		
DRM		19.8	
Homograph		17.0	+2.8% ns
Experiment 4	36		
DRM		9.0	
Homograph		12.3%	−3.3% ns
Experiment 5	64		
DRM		11.2	
Homograph		10.0	+1.2% ns
Weighted average	*214*		
DRM		*12.0*	
Homograph		*11.9*	*+0.1%* ns
Experiment 6[a]	36		
DRM		10.6	
Homograph		4.5	+6.1%*

[a] Relatedness decisions in Experiment 6 were divided by the maximum relatedness judgment possible (i.e., "9") to derive percentage increases for DRM and Homograph critical items.
* $p < .01$.

6-item BAS of Robinson and Roediger was similar to the 12-item BAS in the current study (1.44 vs. 1.43, respectively) and their 3-item BAS was similar to our 6-item BAS (.87 and .72, respectively). Hence, as predicted by associative activation, when the results of the two studies are matched on BAS, rather than number of studied words in a list, our false recall data closely replicate those found by Robinson and Roediger.

The current argument that relational processing influences veridical, but not false memory in the DRM paradigm contrasts with a recent study by Hege and Dodson (2004). In their study, Hege and Dodson told participants to recall both studied items and related items. Using this "inclusion" instruction, they found that recall rates of critical items were lower following picture presentation of lists than following word presentation. They concluded that presenting the items as pictures during study reduced relational processing, which they claimed also reduced false recall. An alternative explanation is that associative activation of the critical item from list items is greater from words than from pictures. This makes sense, given the association norms used to generate the DRM lists were generated from words and not from pictures. It is unclear how strongly activated the critical item would become during the presentation of a pictoral representation of a word associate or how strong of a BAS would be obtained if one conducted a norming study using pictoral representations of the list items as cues. Some authors have indeed argued that semantic (as opposed to associative) relations play a greater role in the identification of pictures than words (Chiarello, 1998; Lupker, 1979, 1988). However, a second possibility is that relational processing does increase false memory, but that such processing is driven by associative, rather than semantic relations. (See below for a discussion of three possible ways associative information could influence false memory.)

The current dissociation between veridical and false memory (with effects of thematic consistency only on veridical memory) joins several past studies in demonstrating that a theoretically important variable (Presentation Duration, Retention Interval, and Dementia) can have differential effects on veridical and false memories. For instance, as with the current experiment, McDermott and Watson (2001) found that increasing the presentation rates of studied items had opposite effects on veridical and false recall. Specifically, veridical recall increased across presentation rates of 20, 250, 1000, 3000, or 5000 ms whereas false recall was highest at 250 ms, but decreased thereafter. Similarly, in our current experiments, veridical recall increased across presentation rates of 80, 200, and 1200 ms whereas false recall decreased between 200 and 1200 ms. In an additional dissociation, several researchers have found that false memories remain relatively stable over retention intervals whereas memory for specific list items decays rapidly (Thapar & McDermott, 2001; Seamon et al., 2002; Toglia et al., 1999). Finally, Balota et al. (1999) found that false recall remains stable across old age and dementia whereas veridical memory for list items drastically declines.

Balota et al. (1999) presented an "attentional control" framework to interpret the performance of older adults and Alzheimer's patients in the DRM paradigm. [This has considerable similarity to the activation monitoring view of Roediger, Balota et al. (2001).] According to this view, participants must exert attentional control to discriminate between (a) highly activated representations of nonpresented critical items and (b) item-specific information produced during study (see Balota et al., Fig. 5). In this light, one might argue that DRM performance has some similarities to the selection process involved in the Stroop task (Stroop, 1935), wherein a participant must control the dominant tendency to

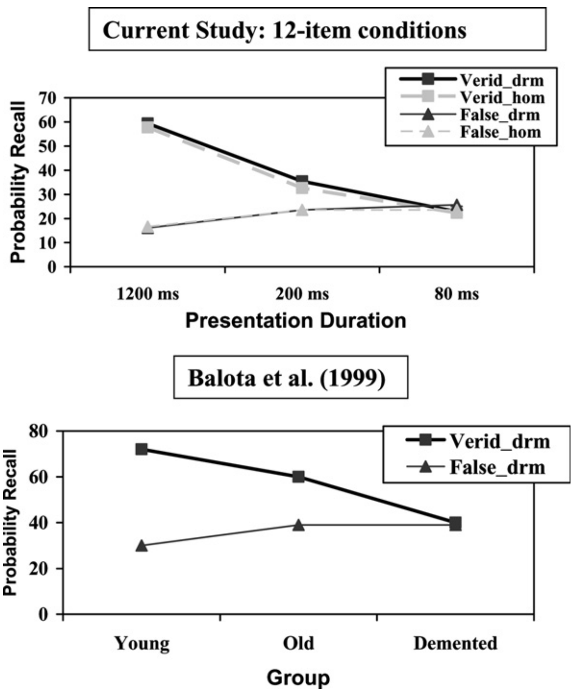

Fig. 5. Comparison of current study to Balota et al. (1999).

respond with the more strongly activated representation. This analogy provides a potential explanation of age and dementia-related changes in false memory performance, as there is evidence of age and dementia-related declines in the ability to inhibit the strongly activated "word" responses in Stroop tasks (Spieler, Balota, & Faust, 1996). Interestingly, Sommers and Huff (2003) have recently found Stroop interference effects to predict false recall for both young and elderly subjects, above and beyond baseline differences in response latency.

The relation between presentation duration effects on young adults in the current experiment and effects of aging and dementia in Balota et al. (1999) are shown in Fig. 5. What is apparent in comparing the two studies is that false memories remain stable across impaired groups (older adults and demented patients) and impaired item-encoding conditions (short presentation) whereas veridical memories show a sharp decline in recall in both cases. The Balota et al. attentional control perspective accounts for both presentation duration effects and retention interval effects if one makes the assumption that item-specific information (a) requires longer presentation durations and (b) fades more quickly over time. The top panel of Fig. 5 also highlights our finding that both 12-item lists that converge on the same meaning and 12-item lists that converge on different meanings produce identical results. Benjamin (2001) provided additional evidence for the influence of both activation and item-specific components.

Although the current results are consistent with many past false memory studies, they are at odds with the past literature on the effects of blocking. In fact, McDermott (1996), Mather, Henkel, and Johnson (1997) and Toglia et al. (1999) have obtained greater false memories when related list items are presented in a "blocked" order, relative to a "random" order. However, these other studies used many more lists (between 45 and 100 items presented) with items randomly intermixed during presentation. In contrast, in the current study we used two 12-item lists (24 items presented) in which all the items were related to the critical item in an alternating manner, allowing for implicit summation of activation for the critical item during presentation. It is quite possible that both widely separating related items during study and presenting several different themes per study list are necessary to produce a blocking effect in false memory. Nonetheless, it should be noted that blocking effects, when present, can be accounted for more easily by meaning-based than by activation-based explanations. Perhaps future research can delineate which methodological changes from the standard presentation conditions (greater number of items or themes per list, longer retention interval, etc...) are necessary to produce blocking effects on false memory.

Methodological concerns

A possible criticism regarding the current methodology for exploring false memory is that the use of different list items and critical items in the DRM and Homograph lists confounds false memory with item differences. Indeed, it would have been preferable to compare the current 12-item mixed homograph (e.g., 6 items relating to the "stumble" meaning and 6 items related to the "season" meaning) list to a homograph list with 12 associates related to only one meaning (e.g., 12 "stumble" items or 12 "season" items). This would have allowed us to examine thematic consistency while holding the critical items constant. Of course, using different critical items in the DRM and Homograph lists does not diminish the strength of the inferences drawn from our design for two major reasons. First, our 6-item condition served as a baseline measure of false memory in which both DRM and Homograph critical nonpresented targets were preceded by 6 associates that all converged on a single meaning. Across all experiments, there was NO difference in false recall between DRM and Homograph items in this baseline condition, suggesting that levels of false recall were indeed matched across the different item types. Second, even if baseline levels of false recall did differ it still would not be a problem because we were not interested in comparing absolute levels of false memory between Homograph and DRM lists, but rather comparing the *relative increase* in false memory a critical item shows as the number of associates increased from 6 to 12 items. However, we do concede that our design precluded us from examining any possible interactions of associative strength and meaning consistency by using a 2×2 design.

A second possible criticism regarding the current methodology is that our manipulation of meaning was too weak to obtain a difference in false memory. As noted previously, our overall levels of false memory in the current experiment were lower than those obtained in earlier experiments. Because of this, one could argue that our design was not sensitive enough to detect effects of thematic consistency. However, the present results consistently yielded effects of thematic consistency. In every experiment we obtained a significant interaction between the number of related items (6 vs. 12) and the item type (DRM vs. Homograph). Adding 6 associates was always more beneficial (or less detrimental) to veridical recall of items from DRM lists than items from Homograph lists. Indeed, across four recall experiments using mixed lists (Experiments 2, 3a, 3b, and 4),veridical recall of DRM items increased $1.9\% \pm 1.3\%$ while Homograph items decreased $5.5\% \pm 1.4\%$. In all experiments, this interaction was significant. In contrast, the increase in false recall was $12.4\% \pm 3.1\%$ for DRM critical items and $12.8\% \pm 2.6\%$ for Homograph critical items and this 0.4% interaction contrast never approached significance [all individual Fs < 1, combined $F(1, 149) = 0.47$]. Thus, relative to the effect of number of related items on veridical recall, the effect of number of related items on false recall was actually quite large. Thus, our design was indeed sensitive enough to pick up even small differences in recall patterns between DRM and Homograph items. Indeed, in Experiment 6, we did obtain a significant interaction between Item Type and Number related for our critical items using a relatedness judgment task. Despite such sensitivity, we consistently failed to obtain any evidence of an interaction between number of related items and item type for false recall.

A third possible criticism is that perhaps one list of 12 words all related to the same meaning (e.g., the SLEEP list) might yield the same "meaning-based" familiarity as two 6-word lists that converge on different meanings (e.g., the "autumn" meaning of FALL plus the "stumble" meaning of FALL). Thus, the equation [$FALL_1 + FALL_2 = SLEEP$] where $FALL_1$ refers to the "autumn" meaning and $FALL_2$ refers to the "slip" meaning. This "additive theme" account could certainly explain our Experiment 1 data in which adding the 6 second-meaning additional associates did not affect veridical recall for homograph list items, yet increased false recall of critical items to the same degree as adding 6 same meaning associates did for DRM critical items. However, the "additive theme" account could not explain how intermixing the meanings of homograph lists selectively disrupted veridical recall of those lists while not influencing false recall. Nor could the additive theme account explain our significant difference in "relatedness judgments" in Experiment 6. Instead, the most plausible explanation for these results is that intermixing the meanings reduced the amount of relational processing during study. Thus, each theme was weaker in the intermixed condition relative to the blocked 6-item condition. Nonetheless, despite this difference, we still found the same increase in false memory. The decrease in "know" responses in the intermixed condition found in Experiment 5 for hit rates only supports the argument that disrupted relational processing hurts veridical memory while not influencing false memory.

Three associative processes and false memories

Implicit associative summation

Balota and Paul (1996) proposed that activation summates over associated primes such that additional primes continue to increase the level of activation of a related target's memory representation. The strongest evidence for this claim was their continued finding that priming effects in their RR condition (e.g., kidney-piano-ORGAN) were equal to the sum of the individual priming effects in the RU and UR conditions (e.g., kidney-soda-ORGAN + wagon-piano-ORGAN, respectively), relative to the UU condition (e.g., wagon-soda-ORGAN). Similarly, in the current study, the false memories we observed may have reflected a buildup of implicit activation for the critical item "fall" during the presentation of the associated study items "stumble, autumn, slip, season, rise, spring, trip, leaves, faint, brisk, clumsy, harvest." As suggested by Hutchison (2003), this implicit priming mechanism is primarily associative, rather than semantic, in nature. The finding of approximately equal patterns of false memory across associative lists that are thematically consistent (i.e., DRM) and diverging (i.e., homograph) supports this claim. As noted earlier, several other researchers have also observed a strong relationship between false memory and BAS in their experiments (Deese, 1959; Hicks & Hancock, 2002; Roediger, Balota et al., 2001; Robinson & Roediger, 1997). Thus, a familiar picture emerges across both word recognition and memory paradigms, one of an implicit activation mechanism that summates from multiple primes to strongly activate an associated target. This mechanism is similar to that proposed by Anderson (1983).

One potential problem with this "implicit associative summation" account, however, is that false memories tend to persist over a longer time period than veridical memories (Seamon et al., 2002; Thapar & McDermott, 2001; Toglia et al., 1999). This should not occur if implicit activation decays over time, as is commonly assumed (Anderson, 1983; Anderson & Pirolli, 1984; Higgins, Bargh, & Lombardi, 1985). Because of this, Gallo and Roediger (2002) concluded that delay effects provide the strongest evidence for an effect of meaning above and beyond those of activation. It is possible that because the present study used immediate, rather than delayed, memory tests associative activation produced the largest effect, as opposed to meaning-based information.

Associative pathway marking during study

Another way of conceptualizing the present results is within the classic Anderson and Bower (1973, 1974) FRAN and HAM models. Within this framework, conceptual representations (i.e., nodes) become activated during learning and become "marked" as having occurred during a particular list context (e.g., the node for the word "rest" becomes marked with list N). These models further proposed that associative pathways connecting from these concepts are also marked in order to possibly connect them to other studied list items (e.g., "rest-SLEEP-dream"). This allows the person to leave his or her own idiosyncratic "marked" trail during study which can then be used to guide them through later recall by searching the marked associative pathways linking list item to list item. Ultimately, this process allows the person to recall many more items than if s (he) simply relied on what could be briefly stored in working memory.

Given the Anderson and Bower (1973, 1974) account, it is easy to understand how a strongly associated critical item could be incorporated into an episodic memory network during study. As the pathways linking words such as "rest" and "dream" become marked, so too could the pathways connecting these words to the critical nonpresented item SLEEP. Over the entire list, multiple activation of pathways connected to SLEEP could lead to this item receiving more support in the episodic network than many of the items actually presented. As time passes (and contextual markers begin to fade), the large number of marked pathways containing SLEEP would make this node more resistant to forgetting than the actual studied items. The present results simply suggest that activation of pathways, such as those connecting "rest" and "dream" to SLEEP, would primarily be driven by associative, rather than semantic, information.

Associative activation during retrieval

Anderson (1983) further argued that people activate associative pathways during memory retrieval to help generate further possible candidates to be either accepted or rejected as having occurred during the previous study task (the so-called "generate-recognize" strategy). According to Anderson, activating the strongest associates of recalled items increases the chance that test associates will overlap with study associates. Through use of retrieval-based association, repeated convergence upon the same critical item could create a sufficiently strong sense of familiarity to justify its status as a studied item. Indeed, this situation could create the need for attentional control in discriminating between strong associate items repeatedly activated during retrieval from those that were actually studied. It is further predicted that this process should become more difficult as time passes and item-specific contextual information fades. Finally, this retrieval view does not require either the assumption that a critical item's activation persists across long delays (e.g., the implicit activation mechanism) or that the item's representation is embedded within a persistent network of marked pathways (e.g., the pathway marking mechanism). Instead the item only needs to become activated (or perhaps re-activated) during the later memory test. As with the other two processes, this associative retrieval process could explain both the importance of BAS in producing false recall and the equivalent patterns of false recall for DRM and homograph critical items when BAS is equated.

It should be noted that these three possible processes are complimentary. All three processes could contribute to false memory, with their separate contributions becoming observable under different circumstances. For instance, under fast presentation conditions, false memories might primarily reflect implicit associative activation (and perhaps associative activation at retrieval if one or two items are perceived consciously). In contrast, at longer retention intervals false memories may reflect either (1) the critical item being highly entrenched within the dense associative network of studied concepts or (2) the activation of the critical item during testing. Finally, these processes could potentially interact with each other. For instance, the probability of a critical item (or the associative pathway containing the critical item's representation) becoming "marked" may depend upon whether it reaches a certain threshold of activation during study. The critical contribution of the present series of experiments is that these three mechanisms appear to reflect associative level information.

Conclusions

To summarize, the current false memory results provide a conceptual replication of the summation of activation found by Balota and Paul (1996) using lexical decision and pronunciation tasks. Across 4 recall experiments using 3 different presentation durations and 1 recognition experiment at the longest duration, there was no evidence for any effect of thematic consistency on false recall. Hence, all experiments supported the hypothesis that false recall primarily reflects associative activation. In every case, false recall increased substantially when adding additional associates during study and, more importantly, this increase was virtually identical regardless of whether the additional items were of a consistent or inconsistent meaning. In contrast, as predicted, veridical recall on list items and relatedness judgments on critical items did show effects of meaning, with increased recall when items converged on the same meaning during presentation and decreased recall when the presentation of two separate meanings was intermixed rather than blocked. Moreover, critical items following studied items that converged on the same meaning during study were rated as more strongly related to the meaning of the studied items. The results suggest a dissociation of the effects of meaning on veridical and false recall. Actively organizing studied material according to meaning enhances one's recall for that information, but does not influence the activation of implicit nonpresented associates. Instead, "memory" for these ideas appears to be mostly driven by the passive process of automatic associative activation.

Appendix A. DRM and homograph lists used in experiments 1–4

DRM Lists						
"Army" List Total BAS = 1.06						
List 1:	Soldier	General	March	Uniform	Defense	Squad
BAS	.28	.09	.04	.04	.04	.04
List 2:	Military	Draft	Reserve	Marines	Patriot	Captain
BAS	.26	.12	.07	.05	.02	.01
"Anger" List Total BAS = 1.36						
List 1:	Insult	Enrage	Revenge	Tantrum	Violent	Vent
BAS	.30	.24	.10	.05	.04	.04
List 2:	Frustration	Temper	Repress	Hate	Argument	Explode
BAS	.23	.27	.04	.03	.02	.01
"Black" List Total BAS = 1.41						
List 1:	Coal	Burnt	Crow	Color	Coffee	Limosine
BAS	.28	.14	.11	.07	.06	.04
List 2:	Gray	Tar	Darkness	Mascara	Funeral	Death
BAS	.36	.15	.11	.05	.03	.01
"Car" List Total BAS = 1.65						
List 1:	Jeep	Gear	Taxi	Highway	Key	Bike
BAS	.24	.14	.12	.11	.11	.10
List 2:	Drive	Steer	License	Chase	Model	Train
BAS	.48	.11	.09	.05	.05	.05
"Dance" List Total BAS = 1.22						
List 1:	Prom	Hustle	Perform	Twist	Modern	Rhythm
BAS	.22	.13	.08	.08	.06	.04
List 2:	Tap	Hula	Song	Formal	Movement	Motion
BAS	.20	.18	.11	.08	.02	.02
"Fruit" List Total BAS = 0.85						
List 1:	Vegetables	Orchard	Mixed	Lemon	Punch	Juice
BAS	.20	.06	.05	.04	.04	.03
List 2:	Strawberry	Forbidden	Tangy	Peel	Cobbler	Blueberry
BAS	.21	.12	.03	.03	.02	.02
"Ghost" List Total BAS = 1.08						
List 1:	Phantom	Superstition	Believe	Demon	Slime	Fright
BAS	.40	.04	.03	.03	.03	.01
List 2:	Haunt	Spirit	Halloween	Holy	Unseen	Skeleton
BAS	.26	.13	.08	.03	.02	.02
"Girl" List Total BAS = 1.15						
List 1:	Gal	Virgin	Dress	Sexy	Beautiful	Sister
BAS	.32	.06	.06	.05	.04	.04
List 2:	Pretty	Doll	Female	Material	Makeup	Friend
BAS	.14	.12	.09	.09	.07	.07

(*continued on next page*)

Appendix A (*continued*)

"Music" List Total BAS = 0.75						
List 1:	Cello	Country	Arts	Noise	Scales	Tour
BAS	.13	.10	.06	.03	.03	.02
List 2:	Piano	Rock	Chamber	Sing	Video	Hear
BAS	.23	.04	.03	.03	.03	.02
"Needle" List Total BAS = 1.21						
List 1:	Pin	Sew	Heroin	Pine	Stitch	Point
BAS	.28	.15	.06	.05	.04	.02
List 2:	Injection	Knitting	Prick	Sharp	Shot	Yarn
BAS	.33	.13	.10	.03	.01	.01
"Paper" List Total BAS = 1.43						
List 1:	Notebook	Copier	Printer	Research	Plastic	Toilet
BAS	.27	.14	.10	.09	.06	.05
List 2:	Typewriter	News	Graph	File	Tear	Receipt
BAS	.24	.17	.15	.07	.05	.04
"Pain" List Total BAS = 0.81						
List 1:	Relief	Sprain	Pressure	Orthodontist	Migraine	Stomach
BAS	.11	.08	.08	.06	.05	.03
List 2:	Injection	Bruise	Harmful	Threshold	Cut	Nerve
BAS	.10	.09	.07	.06	.05	.03
"Sleep" List Total BAS = 2.19						
List 1:	Slumber	Lay	Motel	Trance	Lazy	Nightmare
BAS	.51	.20	.14	.13	.06	.06
List 2:	Snooze	Wake	Bedroom	Unconscious	Deep	Blanket
BAS	.52	.27	.18	.08	.02	.02
"Smell" List Total BAS = 1.33						
List 1:	Sniff	Nose	Sweat	Rank	Rotten	Fish
BAS	.44	.11	.07	.07	.06	.04
List 2:	Perfume	Rose	Fresh	Touch	Soap	Trash
BAS	.39	.03	.04	.03	.02	.03
"Soft" List Total BAS = 1.36						
List 1:	Fluffy	Gentle	Cushion	Comfort	Lamb	Kitten
BAS	.26	.19	.12	.04	.04	.03
List 2:	Tender	Plush	Skin	Furry	Warm	Feather
BAS	.29	.17	.16	.06	.05	.04
"Teeth List" Total BAS = 1.40						
List 1:	Plaque	Jaws	Tongue	Gum	Hygiene	Chew
BAS	.41	.12	.06	.04	.04	.02
List 2:	Fangs	Mouth	Bite	Grit	Nails	Drill
BAS	.31	.18	.09	.09	.02	.02
"Tell" List Total BAS = 1.28						
List 1:	Confess	Instruct	Report	Admit	Story	Secret
BAS	.23	.12	.10	.09	.05	.05
List 2:	Notify	Show	Say	Hint	Advise	Warning
BAS	.39	.12	.04	.04	.03	.02
"Thief" List Total BAS = 0.60						
List 1:	Stolen	Sneaky	Con	Suspect	Beggar	Liar
BAS	.16	.05	.02	.02	.02	.02
List 2:	Burglary	Rob	Crooked	Outlaw	Crime	Jewel
BAS	.11	.07	.05	.05	.02	.01
Homograph lists						
"Bat" List Total BAS = 1.08						
List 1:	Baseball	Ball	Softball	Hit	Club	Glove
BAS	.31	.19	.09	.04	.02	.02

Appendix A (*continued*)

List 2:	Vampire	Blind	Cave	Fangs	Cavern	Rabies*
BAS	.21	.13	.05	.01	.01	.00
"Block" List Total BAS = 0.58						
List 1:	Barrier	Obstacle	Impediment	Shield	Obstruct*	Barricade
BAS	.09	.06	.04	.01	.12	.03
List 2:	Brick	Concrete	Corner	Cube	Cement	Wooden*
BAS	.05	.04	.04	.03	.03	.04
"Bow" List Total BAS = 1.23						
List 1:	Arrow	Indian	Cross	Archery*	Violin	Hunter*
BAS	.53	.03	.02	.24	.03	.00
List 2:	Ribbon	Tie	Sash	Gift	Present	Decoration
BAS	.26	.10	.01	.01	.00	.00
"Cell" List Total BAS = 1.03						
List 1:	Nucleus	Nerve	Molecule	Bacteria	Germ	Membrane*
BAS	.10	.10	.04	.02	.012	.48
List 2:	Prison	Jail	Inmate	Cage	Holding*	Bars*
BAS	.10	.07	.05	.01	.16	.04
"Check" List Total BAS = 0.63						
List 1:	Payment	Bill	Welfare	Deposit	Blank	Cash
BAS	.13	.08	.05	.02	.01	.01
List 2:	Inspect	Examine	Evaluate	Review	Verify*	Overlook*
BAS	.13	.10	.03	.02	.08	.00
"Class" List Total BAS = 0.95						
List 1:	Course	Subject	Lecture	Lesson	Attendance	Professor
BAS	.19	.14	.12	.11	.10	.00
List 2:	Stereotype	Group	Economic	Status	Prestige	Social
BAS	.20	.05	.02	.01	.01	.00
"Date" List Total BAS = 0.78						
List 1:	Prom	Blind	Couple	Double	Relationship	Romantic
BAS	.13	.07	.02	.02	.01	.01
List 2:	Calendar	Appointment	Year	Birth	Event	Time
BAS	.30	.11	.07	.02	.02	.00
"Fall" List Total BAS = 3.20						
List 1:	Stumble	Slip	Rise	Trip	Faint	Clumsy
BAS	.70	.51	.33	.29	.21	.09
List 2:	Autumn	Season	Spring	Leaves	Brisk	Harvest
BAS	.52	.19	.19	.17	.00	.00
"Fly" List Total BAS = 3.10						
List 1:	Swatter	Maggot	Moth	Insect	Bug	Mosquito
BAS	.74	.15	.10	.09	.06	.05
List 2:	Kite	Soar	Airplane	Birds	Wings	Glide
BAS	.57	.37	.32	.32	.23	.10
"Foot" List Total BAS = 2.55						
List 1:	Toe	Ankle	Shoe	Heel	Leg	Sock
BAS	.60	.36	.32	.22	.20	.17
List 2:	Inch	Yard	Mile	Meter	Measurement	Length
BAS	.47	.12	.05	.02	.02	.00
"Head" List Total BAS = 2.11						
List 1:	Chairperson	Chief	Leadership	Department	Boss	Principal
BAS	.13	.05	.04	.04	.03	.02
List 2:	Skull	Hat	Neck	Scalp	Brain	Hair
BAS	.49	.35	.33	.32	.27	.04

(*continued on next page*)

Appendix A (*continued*)

"Horn" List Total BAS = 1.85						
List 1:	Honk	Bugle	Trumpet	Tuba	Trombone	Brass
BAS	.52	.30	.24	.08	.06	.04
List 2:	Unicorn	Ram	Bull	Antler	Antelope	Rhinoceros*
BAS	.11	.11	.05	.05	.03	.24
"Iron" List Total BAS = 1.13						
List 1:	Crease	Starch	Steam	Press	Wrinkle	Shirt
BAS	.09	.09	.09	.06	.06	.00
List 2:	Ore	Steel	Metal	Rust	Copper	Element
BAS	.32	.22	.10	.05	.03	.02
"Letter" List Total BAS = 2.39						
List 1:	Number	Greek	Bold	Symbol	Print	Scarlet*
BAS	.17	.04	.02	.02	.01	.48
List 2:	Envelope	Stamp	Note	Mail	Send	Recommendation
BAS	.49	.38	.29	.24	.15	.11
"Race" List Total BAS = 1.23						
List 1:	Drag	Compete	Track	Runner	Winner	Horse
BAS	.21	.12	.10	.05	.04	.03
List 2:	Minority	Bias	Prejudice	Stereotype	Culture	Ethnicity*
BAS	.06	.06	.04	.04	.01	.48
"Right" List Total BAS = 2.81						
List 1:	Wrong	Correct	Accurate	Proper	Exact	Answer
BAS	.72	.23	.16	.07	.07	.04
List 2:	Left	Starboard	Clockwise	Turn	Direction	Handed*
BAS	.93	.10	.06	.06	.05	.32
"Roll" List Total BAS = 0.75						
List 1:	Biscuit	Bun	Muffin	Crescent	Cinnamon	Bread
BAS	.13	.11	.08	.07	.05	.03
List 2:	Dice	Tumble	Shake	Wheel	Bounce	Rotate*
BAS	.12	.12	.02	.01	.01	.00
"Watch" List Total BAS = 1.65						
List 1:	Observe	Television	Guard	Neighborhood	Look	Patrol
BAS	.50	.12	.11	.08	.03	.03
List 2:	Wrist	Time	Clock	Pocket	Tick	Digital*
BAS	.34	.15	.08	.04	.01	.20

* Indicates cue words not found in Nelson et al. (1999). These association strengths were instead calculated in a norming procedure based on 25 participants conducted in our own laboratory at Washington University.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 261–295.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Oxford, England: Winston and Sons.

Anderson, J. R., & Bower, G. H. (1974). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123.

Anderson, J. R., & Pirolli, P. L. (1984). Spread of activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 791–798.

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explorations from a global matching perspective. *Journal of Memory and Language, 39*, 371–391.

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48*, 1–15.

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

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 336–345.

Balota, D. A., & Paul, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 827–845.

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 941–947.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology, 49*, 229–240.

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. Special issue: Limited resource models of cognitive development. *Developmental Review, 10*(1), 3–47.

Brainerd, C. J., & Reyna, V. F. (1998). When things that were never experienced are easier to "remember" than things that were. *Psychological Science 9* (6) 484–489.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164–169.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331–350.

Chiarello, C. (1998). On codes of meaning and the meaning of codes: Semantic access and retrieval within and between hemispheres. In M. Beeman & C. Chiarello (Eds.), *Right hemisphere language communication: Perspectives from cognitive neuroscience* (pp. 141–160). Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, B. H. (1963). Recall of categorized words lists. *Journal of Experimental Psycology, 66*(3), 227–234.

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

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Crowder, R. G. (1976). *Principles of learning and memory*. Oxford, England: Lawrence Earlbaum.

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

Dodson, C. S., & Schacter, D. L. (2002). When false recognition meets metacognition: The distinctiveness heuristic. *Journal of Memory and Language, 46*, 782–803.

de Morney Davies, P. (1998). Automatic semantic priming: The contribution of lexical- and semantic-level processes. *European Journal of Cognitive Psychology, 10*, 389–412.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gallo, D. A., & Roediger, H. L. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language, 47*(3), 469–497.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory, 10*(2), 83–98.

Hege, A. C. G., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both impoverished relational-encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(12), 787–795.

Higgins, E. T., Bargh, J. A., & Lombardi, W. J. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 59–69.

Hines, D., Czerwinski, M., Sawyer, P. K., & Dwyer, M. (1986). Automatic semantic priming: Effect of category exemplar level and word association level. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 370–379.

Hicks, J. L., & Hancock, T. W. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin & Review, 9*(4), 807–815.

Hutchison, K. A. (2003). Is semantic priming due to association strength or featural overlap? A micro-analytic review. *Psychonomic Bulletin & Review, 10*(4), 785–813.

Jenkins, J. J., & Russell, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology, 47*, 818–821.

Kawamoto, A. H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language, 32*, 474–516.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*, 618–630.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*(6), 485–495.

Lupker, S. J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior, 23*, 709–733.

Lupker, S. J. (1988). Picture naming: An investigation of the nature of categorical priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 444–445.

Mantyla, T. (1997). Recollections of faces: Remembering differences and knowing similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1203–1216.

Marshall, G. R. (1967). Stimulus characteristics contributing to organization in free recall. *Journal of Verbal Learning & Verbal Behavior, 6*(3), 364–374.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 3–23.

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*(6), 826–837.

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language Special Issue: Illusions of memory, 35*(2), 212–230.

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

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27*, 545–559.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Miller, M. B., & Wolford, G. L. (1999). Theoretical commentary: The role of criterion shift in false memory. *Psychological Review, 106*(2), 398–405.

Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science: Journal of Neural Computing, Artificial Intelligence and Cognitive Research, 6*, 413–427.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. (1999). The University of South Florida word association, rhyme and word fragment norms. Available from: http://w3.usf.edu/FreeAssociation.

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.

Plaut, D. C. (1995). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17*, 291–321.

Posner, M. I., & Snyder, C. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Hillsdale, NJ: Lawrence Erlbaum Press.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the prepositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior, 17*(4), 403–417.

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

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger III & J. S. Nairne, et al. (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington, DC: American Psychological Association.

Roediger, H. L., & Gallo, D. A. (2003). Processes affecting accuracy and distortion in memory: An overview. In M. L. Eisen (Ed.), *Memory and suggestibility in the forensic interview* (pp. 3–28). Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*(3), 385–407.

Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language, 43*(1), 67–88.

Schacter, D. L., Cendan, D. L., Dodson, C. S., & Clifford, E. R. (2001). Retrieval conditions and false recognition: Testing the distinctiveness heuristic. *Psychonomic Bulletin & Review, 8*, 827–833.

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories. The effect of retention interval on recall and recognition. *Memory & Cognition, 30*(7), 1054–1064.

Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psycology: Learning, Memory, & Cognition, 18*, 1191–1210.

Sommers, M. S., & Huff, L. M. (2003). The effects of age and dementia of the Alzheimer's type on phonological false memories. *Psychology and Aging, 18*(4), 791–806.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 461–479.

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

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

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*(3), 424–432.

Thompson, C. P., Hamlin, V. J., & Roenker, D. L. (1972). A comment on the role of clustering in free recall. *Journal of Experimental Psychology, 94*, 108–109.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language, 38*, 440–458.

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7*, 233–256.

Tulving, E. (1985). How many memory systems are there?. *American Psychologist, 40*(4), 385–398.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition, 22*(1), 111–126.

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

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

Whittlesea, B. W. A. (2002). False memory and the discrepancy-attribution hypothesis: The prototype-familiarity illusion. *Journal of Experimental Psychology: General, 131*, 96–115.