Output position and word relatedness effects in a DRM paradigm: Support for a dual-retrieval process theory of free recall and false memories

Terrence M. Barnhardt *, Hyun Choi, David R. Gerkens, Steven M. Smith

Department of Psychology, Texas A&M University, College Station, TX 77843, USA

Received 28 November 2005; revision received 24 April 2006

Abstract

Five experiments investigated predictions—derived from a dual-retrieval process approach to free recall (Brainerd, C. J., Wright, R., Reyna, V. F., & Payne, D. G. (2002). Dual-retrieval processes in free and associative recall. Journal of Memory and Language, 46, 120–152.)—about false memories in a DRM-like paradigm. In all the experiments, the presence of the critical words in the study lists was manipulated within subjects. In all the experiments, the output position of presented critical words was earlier than the output position of nonpresented critical words and the output positions of both types of words was closer to the center than to the ends of the recall protocols. In Experiments 2–5, unrelated words were intermixed with related words in the study lists. In all of these experiments, recall of related words was greater than recall of unrelated words. However, in Experiments 4 and 5, the advantage for recall of related words was greater after the critical item was output than before it was output. These findings were consistent with the notions that: (1) there are two successive retrieval processes (direct access of verbatim traces and reconstruction from gist traces) in free recall, (2) items are recalled in ascending order of strength during direct access and descending order of strength during reconstruction from gist, and (3) false memories for words are attributable to reconstruction from gist traces.

© 2006 Elsevier Inc. All rights reserved.

Keywords: False memory; Free recall; Output order; Fuzzy-trace theory; Reconstruction; Cognitive triage

The history of psychology is replete with laboratory examples of memory errors—or false memories—in which subjects claim that they had earlier encountered some stimulus that had not actually been presented or had some experience that had not actually occurred (for review, see Roediger, 1996; Schacter, 1995). Many different paradigms have been used to study false memories, but studying false memories within the context of traditional list-learning experiments has sharply increased since Roediger and McDermott (1995) reintroduced an approach first used by Deese (1959). In what has come to be known as the Deese-Roediger-McDermott (DRM) paradigm, subjects study a series of lists where each list consists of words that are associatively related to a critical nonpresented word. In Roediger and McDermott (1995), subjects performed both free recall and recognition tests. As Roediger and McDermott (1995) pointed out, most experiments concerned...
with memory errors had used recognition tests because such errors had been more reliably observed on this type of test than on free recall tests. For example, in Experiment 2, false recognition of critical nonpresented words that had not been falsely recalled occurred at an even higher rate (65%) than the recognition of presented words that had not been recalled (50%). However, Roediger and McDermott (1995) considered the high rate of false recall the more important finding because reliable observations of robust false recall had previously been so rare. Indeed, in Experiment 1, the false recall rate was 40% (recall rate of studied words was 65%) and in Experiment 2 the false recall rate was 55% (recall rate of studied words presented in the middle of the studied lists was 47%).

In explaining false memory effects of this kind, Roediger and his colleagues (e.g., Robinson & Roediger, 1997; Roediger, Watson, McDermott, & Gallo, 2001) have offered an activation/monitoring framework. In the theory’s simplest form, the role of reality monitoring (Johnson & Raye, 1981) or source monitoring (Johnson, Hashtroudi, & Lindsay, 1993) in the incidence of false memories principally occurs during test. The job of these mechanisms is to discriminate events that actually occurred from events that were only imagined. These mechanisms have failed, for example, when a subject reports having seen a picture of an object when they had only seen the name of the object (e.g., Lane & Zaragoza, 1995).

The role of activation in the incidence of false memories predominantly occurs at study. In the DRM paradigm, when the list of words is presented, the representation of the critical nonpresented word is highly activated as a function of spreading activation mechanisms (e.g., Collins & Loftus, 1975). Such activation can lead to the occurrence of implicit associative responses, or IAR’s (Underwood, 1965). That is, the activation of the nonpresented critical word may be powerful enough that the word is consciously thought of during the study episode. As such, the idea of the word may become associated with the environmental context in which the list was presented, just like the words that were actually presented (but see Seamon et al., 2002 for a counterexample).

A number of observations support the idea that false memories of nonpresented critical words are very similar to veridical memories of presented words. For example, a high degree of confidence usually accompanies false memories (e.g., Payne, Elie, Blackwell, & Neuschatz, 1996; Roediger & McDermott, 1995), false memories are often accompanied by remember judgments (e.g., Gallo, McDermott, Percer, & Roediger, 2001; Payne et al., 1996), subjects are willing to identify the voice in which a critical nonpresented word was “presented” (e.g., Gallo et al., 2001; Hicks & Marsh, 1999; Math, Henkel, & Johnson, 1997; Payne et al., 1996), and priming for critical nonpresented items on implicit memory tests has been observed (e.g., McDermott, 1997; see also McKone & Murphy, 2000; Smith, Gerkins, Pierce, & Choi, 2002; Tse & Neely, 2005). Priming on such tests has often been attributed to perceptual mechanisms (e.g., Schacter, 1990).

However, true and false memories are not necessarily isomorphic. For example, Mather et al. (1997) used a modified memory characteristics questionnaire (Johnson, Nolde, & De Leonardis, 1996) to ascertain the qualitative characteristics of veridical and false memories. Mather et al. (1997) showed that false memories had less auditory detail and less remembered feelings and reactions than memories for presented words. In addition, whereas veridical memory of presented items tends to decline over a delay, false memory of nonpresented critical items remains relatively stable (e.g., Brainerd, Reyna, & Brandse, 1995a; McDermott, 1996; Payne et al., 1996; Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999). Retained false memories are also more likely to be given a remember judgment after a delay (Payne et al., 1996). Finally, when carefully controlling for demand characteristics, small differences have been found between correct recognition and false recognition in participants’ willingness to attribute items to a particular source and in their confidence in doing so (Lampinen, Neuschatz, & Payne, 1999). However, participants were willing to make source attributions and were confident in their false memories for nonpresented critical items quite often and significantly more often than for unrelated lures (Lampinen et al., 1999).

In contrast to activation/monitoring theory, fuzzy trace theory (e.g., Brainerd, Reyna, & Kneer, 1995b) offers a clear distinction between the nature of the memory trace of a presented word and the nature of the memory representation that supports false memories. In fuzzy trace theory, the encoding of presented words results in the creation of verbatim traces, which are item-specific traces that preserve the surface details of the stimulus. The encoding of the presented words also results in the formation of a gist memory, which is an abstraction of the property or properties that the studied words have in common, like the sense of meaning that can be derived from a list of words that are associatively related. In fuzzy trace theory, gist memories serve as the basis upon which false memories are generated at test.

With respect to free recall tests, Brainerd, Wright, Reyna, and Payne (2002) have proposed a dual-retrieval process theory of free recall in which verbatim and gist traces are differentially accessed at test by two distinct retrieval processes: direct access and reconstruction. In direct access of verbatim traces, “participants recall the targets by merely reading out surface information as it... flashes in the mind’s eye, much as an actor would recite words... seen on a script” (Brainerd et al., 2002, p.
Direct access is the faster and more accurate of the two retrieval processes. As a result, it occurs first during free recall, is associated with high confidence, and is virtually errorless. Finally, because direct access is susceptible to output interference, the weakest verbatim traces are output first and the strongest verbatim traces are output last. Brainerd, Reyna, and their colleagues have referred to this last claim, and the data supporting it, as cognitive triage (e.g., Brainerd, Reyna, Howe, & Kevershan, 1991).

In reconstruction from gist traces, candidates for response are generated from the meaning (e.g., tools) that was abstracted from the list of presented words (e.g., saw, wrench, screwdriver, pliers, drill, chisel, file) and sometimes these candidate responses are not from the original list (e.g., hammer). Reconstruction is the slower and less accurate of the two retrieval processes. As a result, it occurs later during free recall and is associated with lower confidence. These response candidates are subjected to a judgment process that utilizes a confidence criterion to accept or reject reconstructed candidates for response. At times, confidence for a candidate that was not on the study list will be high enough to exceed the criterion and a false memory will occur. Finally, reconstruction is not susceptible to output interference, so candidate responses of which the participant is most confident are output first and candidate responses of which the participant is least confident are output last. Note that the output order of the candidate responses associated with reconstruction is the reverse of that for direct access (i.e., strongest to weakest in reconstruction vs. weakest to strongest in direct access).

One straightforward prediction from the dual-retrieval process approach to free recall is that the probability of observing a false memory should increase as output position increases in a recall protocol. This is because false memories are based on reconstruction, which comes on-line after direct access has been exhausted. Indeed, several reports are consistent with this prediction, both when a DRM paradigm is used (McDermott, 1996; Payne et al., 1996; Roediger & McDermott, 1995) and when it is not (e.g., Schwartz, Fisher, & Hebert, 1998; Sommers & Lewis, 1999; for review, see Reyna & Brainerd, 1995).

With a couple of added assumptions, the relatively late output of nonpresented critical words can also be accounted for by the activation/monitoring framework. The first assumption is that presented words in a DRM paradigm have a higher trace strength than nonpresented words because the traces of presented words are formed via both item-specific and relational processing (e.g., Hunt & McDaniel, 1993), whereas the traces of nonpresented words are formed primarily via relational processing only (Roediger, Watson et al., 2001). The second assumption is that the stronger the memory trace, the earlier it is recalled. Thus, at first glance, it seems that the relatively late output position of nonpresented critical words does not strongly differentiate between these theories.

However, it also appears that the two theories make different predictions regarding the relationship between the output positions of nonpresented critical items and presented critical items. Typically, in the DRM paradigm, the recall of critical nonpresented items is compared to the recall of noncritical presented items, although exceptions exist in which memory for nonpresented critical words is compared with memory for presented critical words (e.g., Miller & Wohlford, 1999; Westerberg & Marsolek, 2003). In the present experiments, the approach will be to compare the output position of the critical word across lists in which it is and is not presented.

With some lists that contain the critical item (the “presented” condition) and others that do not (the “nonpresented” condition), it is possible to use the dual-retrieval approach to generate some predictions regarding output order, granting a couple of additional assumptions. The first additional assumption is that a presented critical item should produce the strongest verbatim trace because it can take advantage of both item-specific and relational processes at study. Einstein and Hunt (1980) have argued that relational processing can augment item-specific information, some of which overlaps with the notion of verbatim traces (see also Roediger, Balota, & Watson, 2001). When this assumption is combined with the notion that the output order for verbatim traces progresses from weakest to strongest, this predicts that a presented critical item should be the last item output by direct access retrieval processes. The second additional assumption is that a nonpresented critical item is the strongest reconstructed item and hence the item most likely to exceed response criterion. When this assumption is combined with the notion that the output order for traces reconstructed from gist progresses from strongest to weakest, this predicts that a nonpresented critical item should be the first item output by reconstructive retrieval processes. In short, the fuzzy trace approach predicts: (a) that the output position of a presented critical item should be prior to that of a nonpresented critical item and (b) that the output position of both presented and nonpresented critical items should be closer to the middle of the recall protocol than to the ends of the recall protocol. This hypothesis will be referred to as the output position hypothesis. The two predictions corresponding to this hypothesis will be evaluated in all six of the experiments reported here. Data consistent with these predictions have been recently reported (Brainerd, Payne, Wright, & Reyna, 2003).

A different prediction can be derived from activation/monitoring theory. As noted earlier, to account for the late recall of a nonpresented critical word, relative to
presented words, the most straightforward assumption to add to activation/monitoring theory is that stronger memory traces are output earlier. Given that memory of presented critical words would be the strongest memory traces, this theory predicts that presented critical words should be output more towards the beginning of recall protocol, rather than towards the middle of the protocol. An alternative version of activation/monitoring theory that includes the notion of cognitive triage will be considered in the general discussion.

Confirmation of the output position hypothesis would provide support for the dual-retrieval framework. In turn, this support could serve as the basis upon which to introduce an additional prediction derived from the framework. In essence, dual-retrieval theory predicts that the output of the strongest item (i.e., the critical presented item) serves as a marker signaling the end of direct access retrieval of verbatim traces and that the output of a critical nonpresented item serves as a marker signaling the beginning of reconstructive retrieval from gist traces. To test whether this was the case, words unrelated to the gist of the list were interspersed among, and studied along with, the words related to the gist. Overall, it seemed fairly reasonable to assume that recall of related words would surpass that of unrelated words. However, dual-retrieval theory predicts that the advantage for related word recall should be less in the direct access phase than in the reconstructive phase. This will be termed the relatedness hypothesis. During the direct access phase, verbatim traces of studied words are retrieved. These traces are more perceptual in nature (e.g., Brainerd et al., 2002). As such, their retrieval should be determined more by the extent to which perceptual information was encoded and less by the extent to which meaning information (such as relation to the gist) was encoded. During the reconstructive phase, candidates for response are predominantly generated from the gist meaning of the list of presented words. Because so many more candidates should be generated that are related to the gist than are unrelated to the gist, the number of related words that exceed the judgment criterion for response should far surpass the number of unrelated words.

In contrast, activation/monitoring theory would predict a different pattern of results. Given that presented/related words would produce more highly activated memory traces than presented/unrelated words, and that the most highly activated traces are output first, activation/monitoring theory would predict that there should be relatively large advantage for related word recall early in the output and that this advantage should decrease, rather than increase, as recall progressed.

The relatedness hypothesis will not be evaluated in Experiment 1 because unrelated items were not included in that experiment. The purpose of Experiment 1 was to establish some initial support for the dual-retrieval framework. However, the relatedness hypothesis will be evaluated in all of the remaining experiments.

In short, five experiments were conducted. In Experiment 1, the output position prediction was tested in a standard DRM paradigm. In Experiment 2, the same DRM lists were used as in Experiment 1, but words unrelated to the gist were included with those words. This allowed both hypotheses (output position and relatedness) to be tested. Experiment 3 was a conceptual replication of Experiment 2, but with different materials and a different design. In Experiments 1–3, buffer items were not included. This feature generated an alternative explanation for the relatively central output position of presented critical words, which was investigated in Experiment 4. In Experiments 1–4, the input position of presented critical words was always central. This feature generated another alternative explanation for the relatively central output position of presented critical words that was investigated in Experiment 5.

**Experiment 1**

**Method**

The design of this experiment included one manipulated within-subject factors: presentation of the critical item (presented vs. nonpresented). A second within-subject factor—recall phase—was derived by dividing the recall protocols into those items recalled before the critical item (precritical recall) and those items recalled after the critical item (postcritical recall).

**Participants**

Eighty-nine participants were tested in the experiment. They were all student volunteers who fulfilled a requirement for their introductory psychology course by participating in the experiment.

**Materials**

Eight associative lists were presented to participants. The associative lists were identical to those used by Roediger and McDermott (1995) except that, for half of the lists, the critical word was presented after the fifth word. The chair, doctor, rough, sleep, smell, smoke, sweet, and window lists were used. Each list consisted of either 11 words (when the critical item was presented) or 10 words (when the critical item was missing). Each participant saw four lists in which the critical item was presented (the presented condition) and four lists in which the critical item was not presented (the nonpresented condition). Both the words in each list and the

---

1 The use of unrelated words was inspired by Spence and Holland (1962) and Spence (1964).
lists themselves were presented in a fixed order. The lists were presented in ABBABAAB order, where “A” was nonpresented and “B” was presented for half of the participants. Because the presence of the critical item in a particular list was counterbalanced across subjects, “A” was presented and “B” was nonpresented for the other half of the participants. Often, words in each DRM list are presented in the order of their forward associative strength, with the strongest first and the weakest last (e.g., Roediger et al., 2001). Forward associative strength (FAS) reflects the frequency with which a related word (e.g., bed) is given in response to a critical word (e.g., sleep). Instead, in this experiment, the fixed order of the words in each list was determined randomly with respect to FAS.

Procedure

Participants were tested in groups of 5–20. Each of the stimulus words was presented one at a time for one second on a computer screen with a large font size (font size = 72). After each list presentation, participants performed a 30-s distraction task, in which a series of seven-character strings (letters and numbers) were presented for 1 s each with a 3-s interval between string presentations. Participants counted the number of digits within each string during presentation and recorded the number during the 3-s interval. Right after the number-counting task, they were given the free recall test. This procedure continued until all 8 lists were studied and tested. Participants were instructed to write down only the words from the lists they studied. They were also told that they could write down the words on the answer form in any order they wanted. Subjects were given 3 min for the recall of each list.

Results and discussion

The significance level for all statistical tests was set at $p < .05$, two-tailed.

Of the 89 participants tested, 78 (88%) accurately recalled a presented critical word at least once and falsely recalled a nonpresented critical word at least once. Across the 78 participants, there were 312 opportunities (78 participants $\times$ 4 lists for each participant) to recall presented critical words and 312 opportunities to recall nonpresented critical words. Participants recalled presented critical words 275 times, or an average of 3.5 times out of every 4 lists (88%). Average recall of related words in the presented condition was 71%. Participants recalled nonpresented critical words 179 times, or an average of 2.3 times out of every 4 lists (57%). Average recall of related words in the nonpresented condition was 68%. In short, the false recall rate in Experiment 1 was quite consistent with other reports in the false memory literature (e.g., Roediger & McDermott, 1995).

The rate of intrusions was quite low, averaging one per approximately every five lists ($M = .17$ per list). As a result of this low rate of occurrence, intrusions in this and the other experiments reported here were essentially ignored.

Critical word output position as a function of presentation condition and mean recall as a function of recall phase and presentation is displayed in Table 1. The aver-

<table>
<thead>
<tr>
<th>Critical word presentation</th>
<th>Precritical mean recall</th>
<th>Critical word output position</th>
<th>Postcritical mean recall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Related</td>
<td>Unrelated</td>
<td>Related</td>
</tr>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented</td>
<td>2.9 (.15)</td>
<td>—</td>
<td>3.9 (.15)</td>
</tr>
<tr>
<td>Nonpresented</td>
<td>4.0 (.20)</td>
<td>—</td>
<td>5.1 (.21)</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented</td>
<td>2.3 (.14)</td>
<td>1.0 (.10)</td>
<td>4.4 (.22)</td>
</tr>
<tr>
<td>Nonpresented</td>
<td>2.6 (.18)</td>
<td>1.4 (.13)</td>
<td>5.1 (.28)</td>
</tr>
<tr>
<td><strong>Experiment 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented</td>
<td>1.9 (.10)</td>
<td>.9 (.08)</td>
<td>3.8 (.16)</td>
</tr>
<tr>
<td>Nonpresented</td>
<td>2.8 (.20)</td>
<td>1.2 (.15)</td>
<td>5.2 (.23)</td>
</tr>
<tr>
<td><strong>Experiment 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented</td>
<td>1.5 (.09)</td>
<td>.9 (.07)</td>
<td>3.3 (.13)</td>
</tr>
<tr>
<td>Nonpresented</td>
<td>2.1 (.15)</td>
<td>1.1 (.11)</td>
<td>4.2 (.24)</td>
</tr>
<tr>
<td><strong>Experiment 5</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented</td>
<td>1.9 (.13)</td>
<td>1.2 (.13)</td>
<td>4.3 (.25)</td>
</tr>
<tr>
<td>Nonpresented</td>
<td>3.0 (.27)</td>
<td>1.7 (.22)</td>
<td>5.8 (.48)</td>
</tr>
</tbody>
</table>

Note. Words unrelated to the gist of the list were not presented in Experiment 1. The standard error for each mean is in parentheses.
age absolute output position of presented critical words was 3.9 and the average absolute output position of nonpresented critical words was 5.1. This difference was significant, \( t(77) = -4.94, \ SEM = .232; \) 95% CI = 7–1.6. The output position of false memories is often reported to be later than that reported here (Payne et al., 1996; Roediger & McDermott, 1995; Schwartz et al., 1998). In many of those reports, however, presentation of a DRM list was followed immediately by the recall test. It is well known that the first items output in an immediate free recall test are often those that are still in short-term memory (e.g., Huang, 1986). If this were the case in those experiments, initial recall from short-term memory could push the output of the false memory to a later position than might be the case if a delay intervened between the end of the list and the beginning of recall. McDermott (1996) also reported that false memories were output a bit earlier when recall was delayed, as was the case here.

Of course, absolute output position will vary as a function of total recall. For example, the last word in a long recall protocol will have a later absolute output position than the last word in a short recall protocol. In Experiment 1, average total correct recall (including critical word, whether presented or not) was 8.1 in the presented condition and 7.8 in the nonpresented condition. This difference was significant, \( t(77) = 2.13, \ SEM = .116. \) For this reason, relative output position was also examined. This was calculated for each participant by simply taking the absolute output position and dividing it by the total number of words recalled. In some respects, relative output position may give a more accurate picture of differences between the presented and nonpresented groups because it adjusts for the total recall in each individual recall protocol. The average relative output position of a presented critical word was .49 and the average relative output position of a nonpresented critical word was .65. This difference was also significant, \( t(77) = -5.59, \ SEM = .028. \)

Mean relative output position can also be used to evaluate a couple of other predictions derived from the dual-retrieval approach that were described in the introduction. First, the mean output position of the critical presented items should be closer to the middle position than to the first position in a recall protocol. Second, the mean output position of the critical nonpresented items should be closer to the middle position than to the last position. The values of the mean relative output positions supported both of these predictions: clearly, .49 is much closer to the center of the recall protocol than to the beginning and .65 is much closer to the center than to the end.

Another indication of the difference between the average absolute output positions of presented critical words and nonpresented critical words was the interaction of presentation and recall phase. If the output position of presented critical words is earlier than that of nonpresented critical words, and the total amount of recall is similar in the two presentation conditions, then it follows that precritical recall should be less, and postcritical recall should be greater, in the presented condition than in the nonpresented condition. This was the pattern observed in this experiment, \( F(1,77) = 29.53, \ MSE = 3.909. \) In short, there was strong evidence in support of the output position hypothesis.

**Experiment 2**

In the previous experiment, support was obtained for the output position hypothesis and, by extension, the dual-retrieval framework. First, the output position of false memories for nonpresented critical items was later than the output position of veridical memories for presented critical items. Second, the output positions of both items were relatively close to the center of the recall protocol. These findings lent support to the idea that the output of presented critical words marks the end of the direct access phase, while the output of nonpresented critical words marks the beginning of the reconstructive phase. In turn, the possibility that the output of presented and nonpresented critical words provides reliable markers of the boundary of these two recall phases made it possible to test the relatedness hypothesis.

In Experiment 2, words unrelated to the gist of the list were included in the study lists. The dual-retrieval theory of free recall states that there are two retrieval mechanisms in free recall that occur in succession: the first is the direct access of verbatim traces and the second is reconstruction from gist traces. The relatedness prediction derived from dual-retrieval theory states that the difference between recall of related and unrelated words should be greater after the output of a critical word than before the output of a critical word. In addition to the relatedness hypothesis, the output position hypothesis was also evaluated in this experiment.

**Method**

The design of this experiment included two manipulated within-subject factors: presentation of the critical item (presented vs. nonpresented) and relatedness (related vs. unrelated) of the study list items to the critical word. To evaluate the relatedness hypothesis, a third within-subject factor—recall phase—was derived by dividing the recall protocols into those items recalled before the critical item (precritical recall) and those items recalled after the critical item (postcritical recall).

**Participants**

One-hundred-and-one participants were tested in the experiment. They were all student volunteers who
fulfilled a requirement for their introductory psychology course by participating in the experiment.

**Materials and procedure**

To construct the lists with both related and unrelated items for this experiment, the same eight associative lists of 10 items each were used as in Experiment 1. Then, 10 words unrelated to the critical word were randomly interspersed among the related words in composing each study list. The only criteria for selecting the 10 unrelated words was that they were unrelated to the gist of the list and, for the most part, were also unrelated across lists. The fact that related and unrelated words were not counterbalanced was addressed in Experiment 3. In the end, there were eight lists, four with 21 words if the list was in the presented condition (10 related, 10 unrelated, and 1 critical) and four with 20 words if the list was in the nonpresented condition (10 related and 10 unrelated). When the critical word was presented, it was the 11th item on the list. No more than three related or three unrelated words were presented consecutively. All other details of the materials and procedure were the same as those in Experiment 1.

**Results and discussion**

Of the 101 participants tested, 76 (75%) accurately recalled a presented critical word at least once and falsely recalled a nonpresented critical word at least once. Across the 76 participants, there were 304 opportunities to recall presented critical words and 304 opportunities to recall nonpresented critical words. Participants recalled presented critical words 245 times, or an average of 3.22 times out of every 4 lists (81%). In the presented condition, average recall of related words was 55% and average recall of unrelated words was 26%. Participants recalled nonpresented critical words 167 times, or an average of 2.2 times out of every 4 lists (55%). In the nonpresented condition, average recall of related words was 54% and average recall of unrelated words was 27%. As noted in the introduction, a false recall rate that is as high as the rate of veridical recall of presented related words has been previously reported (e.g., Roediger & McDermott, 1995).

Critical word output position as a function of presentation condition and mean recall as a function of relatedness, recall phase, and presentation is displayed in Table 1. Experiment 2 again provided support for the output position hypothesis. The average absolute output position of a presented critical word was 4.4 and the average absolute output position of a nonpresented critical word was 5.1. This difference was significant, \( t(75) = -2.61, SEM = .291; 95\% CI = .2–1.3 \). As described in Experiment 1, another indication of the difference between the absolute output positions of presented and nonpresented critical words would be the interaction of presentation and recall phase. Once again, precritical recall was less, and postcritical recall greater, in the presented condition than in the nonpresented condition, \( F(1,75) = 6.85, MSE = 3.092 \). There was little difference between total recall in the presented and nonpresented conditions (9.2 vs. 9.1), \( F < 1 \). When the difference between relative output position in the presented and nonpresented conditions (.48 vs. .57) was examined, this difference was also significant, \( t(75) = -2.67, SEM = .033 \). Again, visual inspection of the mean relative output positions for the presented and nonpresented critical words showed that both were much closer to the middle, than to the ends, of the recall protocol.

This experiment also afforded a test of the relatedness hypothesis, which predicted that the advantage for recall of related words over unrelated words should be greater in the postcritical recall phase than in the precritical recall phase. The relatedness hypothesis was not supported. As can be gleaned from Table 1, when the data were collapsed across presentation, there was a large difference between recall of related (2.5) and unrelated (1.2) words during the direct access phase, \( F(1,75) = 141.86, MSE = .886 \). There was a numerically larger difference between recall of related (3.0) and unrelated (1.5) words during the reconstructive phase, \( F(1,75) = 143.00, MSE = 1.188 \). However, the interaction between recall phase and relatedness was not significant, \( F(1,75) = 1.16, p = .29 \).

In addition, neither the interaction of presented and relatedness or the 3-way interaction were significant, both \( F's < 1.5 \). The main effect of recall phase was significant, \( F(1,75) = 5.92, MSE = 4.24 \). Postcritical recall (4.4) was greater than precritical recall (3.6). Of course, the main effect of relatedness was also significant, \( F(1,75) = 462.76, MSE = .635 \). Participants recalled more than twice as many related words (5.4) as unrelated words (2.6).

In summary, two hypotheses were investigated. First, the output position hypothesis was supported. In the dual-retrieval approach to free recall, direct access of verbatim traces precedes reconstructive retrieval from gist memories. If veridical memories of presented words are attributable to direct access of verbatim traces and false memories of nonpresented words are attributable to reconstructive retrieval, then the latter should have a later output position than the former. This was the result observed here.

Second, the relatedness hypothesis was not supported. Fuzzy trace theory describes verbatim traces as consisting primarily of surface information. This led to the prediction that such traces would be independent of meaning. As a result, retrieval of related and unrelated words during the direct access phase should not differ as greatly as during the reconstructive phase. Reconstructive retrieval is described as being heavily
dependent on meaning and, for this reason, there should be a relatively large difference between retrieval of related and unrelated words during this recall phase. In short, the difference between recall of related and unrelated words should be less before recall of a critical word than after recall of a critical word. Instead, there was a small, but nonsignificant, difference in the advantage for related word recall during the precritical (1.3) and postcritical portions (1.5) of the recall protocols.

Although it was possible to develop some post hoc explanations for this result, there was one obvious methodological possibility. In Experiment 2, the related and unrelated words were different words that could have differed in their memorability. In Experiment 3, the related and unrelated words were counterbalanced to rule out this possibility.

Experiment 3

In this experiment, related and unrelated words switched the nature of their relationship to the critical word across different subjects, so that every word was equally often in the related and unrelated condition. This was accomplished by selecting the 20 DRM lists (each of 10 words length; taken from McDermott & Watson, 2001) that have the highest known propensity for producing false memories (Roediger et al., 2001). These lists were then rank-ordered on this dimension and every other list was assigned to one of two groups of 10 lists. One group of 10 lists served as the DRM lists and the other group of 10 lists served as the unrelated words. To select the 10 unrelated words in each of the 10 DRM lists, one word was taken from each of the 10 lists in the other group. To counterbalance the words across the related and unrelated conditions, the process was reversed. That is, this second group of 10 lists became the tests lists and, to get the unrelated words that would be mixed in with a particular DRM list, one word was taken from each of the lists in the first group. In this way, any differences in the recall of related and unrelated words could not be attributed to the nature of the stimuli themselves.

The counterbalancing of related and unrelated words in Experiment 3 remedied one problem, but it introduced another. That is, it seemed possible that recall of “unrelated” words may be inflated because, even though these words were unrelated to other words within a list, they were related to other words across the lists. Processing of these cross-list relationships could build across the lists, thus improving recall of the unrelated words, relative to the recall of unrelated words in Experiment 2. In truth, we had anticipated the relative strengths and weaknesses of the designs in Experiments 2 and 3. In Experiment 2, the problem of cross-list relatedness (introduced in Experiment 3) was remedied by using words that were not related to the list in which they were presented and were not related to words in the other lists. Of course, as we’ve already pointed out, this meant that the stimuli could not be counterbalanced across conditions in Experiment 2. In the end, it was hoped that the pattern of results across Experiments 2 and 3 would converge on a meaningful conclusion regarding the relatedness hypothesis. In addition, the output position hypothesis was also evaluated in this experiment.

Method

The design of this experiment again included two manipulated within-subject factors: presentation of the critical item (presented vs. nonpresented) and relatedness (related vs. unrelated) of the study list items to the critical word. A third within-subjects factor—recall phase—was derived by dividing the recall protocols into those items recalled before the critical item (precritical) and after the critical item (postcritical).

Participants

Seventy-six participants were tested in the experiment. They were all student volunteers who fulfilled a requirement for their introductory psychology course by participating in the experiment.

Materials and procedure

All details of the materials and procedure were the same as those in Experiment 2, except for the changes already described in the introduction to this experiment. In addition to the eight DRM lists used in the previous two experiments, the twelve lists used were the city, cold, cup, foot, mountain, needle, river, rough, slow, soft, spider, and trash lists.

Results and discussion

Of the 76 participants tested, 57 (75%) accurately recalled a presented critical word at least once and falsely recalled a nonpresented critical word at least once. Across the 57 participants, there were 285 opportunities to recall presented critical words and 285 opportunities to recall nonpresented critical words. Participants recalled presented critical words 201 times, or an average of 3.5 times out of every 5 lists (70%). In the presented condition, average recall of related words was 49% and average recall of unrelated words was 24%. Participants recalled nonpresented critical words 109 times, or an average of 1.9 times out of every 5 lists (38%). In the nonpresented condition, average recall of related words was 47% and average recall of unrelated words was 22%. Recall of presented critical items in this experiment (70%) was substantially less than in Experiments 1 (88%) and 2 (81%). The false recall rate in this
experiment (38%) also was substantially less than in Experiments 1 (57%) and 2 (50%). Both of these differences were attributable to the use of some DRM lists in this experiment that had critical words, as shown in previous norming studies (McDermott & Watson, 2001), that were not as highly associated to their respective related words as in the previous experiments. It is also interesting to note that recall of unrelated words in this experiment (around 23%) was slightly less than recall of unrelated words in Experiment 2 (27%), despite the fact that the unrelated words in this experiment were related across the lists.

Critical word output position as a function of presentation condition and mean recall as a function of relatedness, recall phase, and presentation is displayed in Table 1. Experiment 3 provided strong support for the output position hypothesis. The average absolute output position of presented critical words was 3.8 and the average absolute output position of nonpresented critical words was 5.2. This difference was significant, \( t(56) = -4.50, SEM = .310; 95\% CI = -.82.0 \). As in the previous two experiments, precritical recall was less, and postcritical recall greater, in the presented condition than in the nonpresented condition, producing a significant 2-way interaction of presentation and recall phase, \( F(1,56) = 28.42, MSE = 1.934 \). The difference between total recall in the presented (8.3) and nonpresented (7.8) conditions approached significance, \( t(56) = 1.95, SEM = .244, p = .06 \). When the difference between relative output position (see Experiment 2 for the manner in which this was calculated) in the presented (.45) and nonpresented (.63) conditions was examined, this difference was also significant, \( t(56) = -5.05, SEM = .035 \). Visual inspection of the relative output positions revealed that the output position for the critical presented item was much closer to the center of the recall protocol than to the beginning and that the output position for the critical nonpresented item was much closer to the center than to the end of the recall protocol.

This experiment also afforded a test of the relatedness hypothesis. As in Experiment 2, the relatedness hypothesis was not supported. As can be gleaned from Table 1, when the data were collapsed across presentation, there was a large difference between recall of related (2.4) and unrelated (1.1) words during the direct access phase, \( F(1,56) = 139.15, MSE = .702 \). There was also a large difference between recall of related (2.4) and unrelated (1.2) words during the reconstructive phase, \( F(1,56) = 77.21, MSE = 1.00 \). The interaction of recall phase and relatedness was not significant, \( F(1,56) < 1 \).

The 3-way interaction of presentation, relatedness, and recall phase was also significant, \( F(1,56) = 8.84, MSE = .950 \). This was attributable to the fact that the simple 2-way interaction of related and recall phase in the presented condition was reversed in the nonpresented condition. To be specific, in the presented condition, the advantage for recall of related words in the precritical phase was smaller than in the postcritical phase (1.0 vs. 1.4). In contrast, in the nonpresented condition, the advantage for recall of related words in the precritical phase was greater than in the postcritical phase (1.6 vs. .9). This 3-way interaction fell out of the fact that (a) the recall protocol was divided into two phases around recall of the critical word, (b) the output position of presented critical words was prior to that for nonpresented critical words, and (c) recall of the critical word could not be counted as part of either phase. Counting recall of a presented critical word toward the direct access phase or recall of a nonpresented critical word toward the reconstructive phase would produce a tautology. One of the main purposes of this series of experiments was to gather evidence on whether veridical and false memories were isomorphic. Counting recall of a presented critical word toward the direct access phase and recall of a nonpresented critical word toward the reconstructive phase would be tantamount to deciding, a priori, that they were not isomorphic. As a result, if (a) it is accepted that recall of a presented critical word cannot be counted toward recall of related words in the precritical phase and (b) recall of a presented critical word is early, as it was in this experiment, then (c) it almost necessarily follows that recall of related words in the precritical phase would be reduced relative to what it could have been had recall of the presented critical word been counted toward related word recall. The same type of argument can be applied to recall of related words in the postcritical phase in the nonpresented condition. In short, the 3-way interaction was an artifact of not counting recall of the critical word as part of either the direct access or reconstructive phases.

Neither the main effect of recall phase nor the interaction of presented and relatedness were significant, both \( Fs < 1 \). Of course, the main effect of relatedness was again significant, \( F(1,56) = 267.78, MSE = .651 \). Participants recalled more than twice as many related words (4.8) as unrelated words (2.3).

In summary, Experiment 3 again provided tests of the output position and relatedness hypotheses. As in Experiments 1 and 2, both predictions of the output position hypothesis were well-supported. However, as in Experiment 2, the relatedness hypothesis was not supported. The designs of those two experiments had relative strengths and weaknesses. In Experiment 2, truly neutral words were used, but those words could not be counterbalanced with the related words. Therefore, it was unclear whether the differences in recall performance for related and unrelated words were due to the difference in their relatedness to other words in the list or to a difference in the memorability of the words. In Experiment 3, the related and unrelated words were counterbalanced, but that meant that even though the unrelated words were unrelated to other words...
within a list, they were related to other words across the lists. Despite these differences in the two experiments, the outcome was the same: the related word advantage in recall was not significantly larger after a critical word was output than before. This hypothesis will continue to be evaluated, under slightly different circumstances, in the remaining two experiments.

**Experiment 4**

As just noted, the output position hypothesis has been well-supported in all three of the experiments presented so far. In the introduction, it was noted that both activation/monitoring theory and dual-retrieval theory would predict that the output position of critical presented words would be earlier than that for critical nonpresented words. The prediction that differentiated the two theories was that activation/monitoring theory predicts that the output position of a critical presented word should be nearer the beginning of the recall protocol because it has a very high level of activation. The fact that the output position of critical presented words has been much closer to the center of the protocol in all three experiments thus far has been supportive of the dual-retrieval approach. In this approach, direct access of verbatim traces occurs first during free recall and, during that phase, the items more at risk of not being produced (because of output interference) are output first, while those that are less at risk for interference are output later (i.e., cognitive triage).

However, there is an alternative explanation for the fact that recall of critical presented words was much closer to the center of the protocol than to the beginning. In all three of the experiments presented so far, items in the primacy and recency positions of the studied lists were scored just like the items in the other serial positions. There are a few reasons to think that the primacy and recency portions of the study lists may be the first items output during a free recall protocol. First, Nairne, Riegler, and Serra (1991) have argued that the output order of items at test recapitulates the temporal serial order with which these items are input at study (see also Mandler & Dean, 1969). This would mean that the primacy portions of the lists would be output first. Second, Huang (1986) used a successive single-trial free recall design like that used here and found that the recency portions of lists presented later in the series were often output early in the recall protocol.

In short, the fact that the output position of recalled critical presented items was much nearer the center than the beginning of the recall protocol in Experiments 1–3 may be due to the early output of items in the primacy and recency positions of the lists rather than to the early output of items that are more susceptible to output interference (i.e., cognitive triage). This possibility was investigated in Experiment 4. Six unrelated items—three primacy buffers and three recency buffers—were added to the same DRM lists that had been used in Experiment 2. Even though these buffer items could be output during recall of a list, they were not scored as recalled words. If the central output position of presented critical words was due to the relatively early output of primacy and recency items, and activation-monitoring theory is correct that the most strongly activated items are output first, then the removal of the primacy and recency items from the scoring should move the output position of presented critical words much closer to the beginning of the recall protocol. On the other hand, the dual-retrieval approach would be supported if output of the presented critical words remained closer to the center of the recall protocols.

Like Experiments 1–3, Experiment 4 also compared veridical recall to false recall and allowed for another test of the relatedness hypothesis. With regard to the relatedness hypothesis, Experiment 4 was like Experiment 3 in that the unrelated words were unrelated to other words in the list in which they were presented, but were related across lists. However, there were two differences between this experiment and Experiment 3. First, the unrelated words were not counterbalanced with the related words across subjects, making this experiment like Experiment 2 in that respect. Second, the across-list relationship between unrelated words was categorical, rather than associative, in nature.

Finally, this experiment was actually run as two experiments, both with 69 participants. The purpose of the second experiment was to replicate the first experiment. The only difference in the methods of the two experiments was that slightly different materials were used. To be more specific, the same critical words were used in both experiments, but 36 (45%) of the 80 words related to the critical word (8 lists × 10 related words in each list) were replaced with new words in the second experiment. This was done by using the Nelson, McEvoy, and Schreiber (1999) word association norms. In the end, both forward associative strength—the frequency with which a related word (e.g., bed) is given in response to a critical word (e.g., sleep)—and backward associative strength—the frequency with which a critical word (e.g., sleep) is given in response to a related word (e.g., bed)—were held constant across the two experiments. All results replicated across the two experiments. So, to make better use of expositional space, the two experiments were collapsed into the single experiment reported here.

**Method**

The design of this experiment included the same three within-subject factors that had been manipulated in the two previous experiments: presentation of the critical...
item (presented vs. nonpresented), relatedness (related vs. unrelated) of the study list items to the critical word, and recall phase (precritical vs. postcritical).

Participants
One-hundred-and-thirty-eight participants were tested in the experiment. They were all student volunteers who fulfilled a requirement for their introductory psychology course by participating in the experiment.

Materials and procedure
The same eight associative lists of 10 items each were used as had been used in Experiments 1 and 2. To select the unrelated words, eight words were chosen from each of 16 different categories in the Battig and Montague (1969) norms. These 16 categories were not semantically related to any of the DRM lists. Then one of the eight words from each of the 16 categories (for a total of 16 words) was added to each DRM list. Three words were randomly selected to serve as primacy buffers and three as recency buffers. In the end, there were 8 lists, each with 27 words if the list was in the presented condition—10 related, 10 unrelated (and scored), 6 unrelated (and unscored) and 1 critical—or 26 words if the list was in the nonpresented condition (same as in the presented condition, but with no critical word). When the critical word was presented, it was the 13th item on the list. No more than three related or three unrelated words were presented consecutively. All other details of the materials and procedure were the same as those in Experiments 1–3.

Results and discussion
Of the 138 participants tested, 104 (75%) accurately recalled a presented critical word at least once and falsely recalled a nonpresented critical word at least once. Across the 104 participants, there were 416 opportunities to recall presented critical words and 416 opportunities to recall nonpresented critical words. Participants recalled presented critical words 337 times, or an average of 3.2 times out of every 4 lists (81%). In the presented condition, average recall of related words was 46% and average recall of unrelated words was 25%. Participants recalled nonpresented critical words 182 times, or an average of 1.8 times out of every 4 lists (44%). This was a bit more than in Experiment 3 (38%), but substantially less than in Experiments 1 (57%) and 2 (50%). This may have been due to the presence of the unrelated items. It has been previously observed that the frequency of false memories is less after study of a relatively long list consisting of several DRM lists that are randomly mixed during study than after DRM lists that are presented in a blocked fashion (e.g., Mather et al., 1997; McDermott, 1996; Toglia et al., 1999; Tussing & Greene, 1997). In the nonpresented condition, average recall of related words was 44% and average recall of unrelated words was 22%.

Critical word output position as a function of presentation condition and mean recall as a function of relatedness, recall phase, and presentation is displayed in Table 1. Once again, there was strong support for the output position hypothesis. The average absolute output position of presented critical words was 3.3 and the average absolute output position of nonpresented critical words was 4.2. This difference was significant, \( t(103) = -3.39, \quad SEM = .253; \quad 95\% \ CI = .4–1.4. \) As in the previous three experiments, precritical recall was less, and postcritical recall greater, in the presented condition than in the nonpresented condition, producing a significant 2-way interaction of presentation and recall phase, \( t(1,103) = 21.53, \quad MSE = 2.931. \) The difference between total recall in the presented and nonpresented conditions (8.1 vs. 7.7) was significant, \( t(103) = 2.58, \quad SEM = .169. \) When the difference between relative output position in the presented and nonpresented conditions (.42 vs. .55) was examined, this difference was significant, \( t(103) = -4.31, \quad SEM = .029. \)

Visual inspection of both the absolute and relative output positions of the critical presented words revealed the same pattern as in the previous experiments. On average, both presented critical words and nonpresented critical words were output closer to the center, than to the ends, of the recall protocol. With regard to the output position of presented critical words, it is important to note that this occurred despite the fact that the study lists used in this experiment included recency and primacy buffers that were not scored as correct. In the introduction to this experiment, it was argued that the relatively central output position of critical presented items in the previous experiments could have been due to early output of the primacy and recency portions of the study lists used in those experiments. In this experiment, output of those items was not counted in the output position data. Yet, the average relative output position of presented critical items was not appreciably earlier in this experiment (.42) than in the previous experiments (.49, .46, and .45). This finding was inconsistent with the idea that the relatively central output position for presented critical words was due to the early output of the primacy and recency portions of the study. Instead, the observation of a central output position for presented critical words in this experiment and in all the previous experiments was consistent with the notion of cognitive triage and inconsistent with the somewhat standard notion that the strongest or most highly activated memory is output first.

Unlike the previous two experiments, the relatedness hypothesis was supported in this experiment. Collapsing across presentation, the advantage for related recall over unrelated recall in the postcritical phase (1.4) was greater than that same advantage in the precritical phase.
but a relatedness effect was once again observed. Experiment 5, where buffer items were not included, interpretation will be undermined by the results of the rest of the list. However, it should be noted that this word advantage may have been even greater in the primary and recency portions of the study list than in the previous experiments. Furthermore, it appears that the reduction in this advantage was primarily due to a decrease in the precritical recall of related words. One of the main differences between this and the previous experiments was that this experiment included primacy and recency buffers. Experiments 2 and 3 did not. The primacy and recency portions of the study lists in Experiments 2 and 3 included related words. Since removal of those portions of the lists in Experiment 4 decreased the related word advantage, this suggests that the related word advantage may have been even greater in the primacy and recency portions of the study list than in the rest of the list. However, it should be noted that this interpretation will be undermined by the results of Experiment 5, where buffer items were not included, but a relatedness effect was once again observed.

The 3-way interaction was also significant, $F(1, 103) = 7.66$, $MSE = .848$. Again, this interaction appeared to be an artifact of not counting recall of the critical word as part of either the direct access or reconstructive phase. The 2-way interaction of presented and relatedness was not significant, $F < 1$. The main effect of relatedness was once again significant, $F(1, 103) = 319.83$, $MSE = .755$. Participants recalled nearly twice as many related words (4.5) as unrelated words (2.4). Finally, participants recalled more words after recalling the critical word (4.1) than prior to recalling the critical word (2.8), $F(1, 103) = 31.97$, $MSE = 2.69$.

**Experiment 5**

One of the main purposes of Experiments 4 was to test an alternative explanation for the fact that presented critical words were output near the center of the recall protocol. This finding was consistent with the notion of cognitive triage and, by extension, supportive of the dual-retrieval theory of free recall. It was inconsistent with activation/monitoring theory, which predicted that output of presented critical words should occur relatively early in a recall protocol. However, as explained in the introduction to Experiment 4, the central output of presented critical words in Experiments 1–3 was also consistent with another possibility. That is, it could have been the result of early output of the scored primacy and recency portions of the study list, which pushed back the output position of presented critical words to their more central position. This explanation was tested in Experiment 4 by including primacy and recency buffer items in the study list that were not scored in the recall protocol. If those items were indeed responsible for the central output of presented critical words, their removal should have resulted in a much earlier output position for presented critical words. However, the output position of presented critical words in Experiment 4 was again relatively central in the recall protocols.

The main purpose of Experiment 5 was to test another alternative explanation for the relatively central output of presented critical words. In all five of the previous experiments, the input position of the presented critical word was in the middle of the study list. As noted earlier, Nairne et al. (1991) have argued that items are recalled along their temporal serial order. From this point of view, the relatively central output position of presented critical words may be a function of it’s central input position, rather than cognitive triage. In Experiment 5, this idea was tested by varying the input position of the presented critical words. Varying the input position of presented critical words should not alter the central output position of those words if their output position is attributable to cognitive triage.

As in previous experiments, the output position hypothesis for the comparison between presented and nonpresented critical words was tested, as was the relatedness hypotheses.

**Method**

As in Experiments 2–4, the design of this experiment included three within-subjects manipulations. They were presentation of the critical item (presented vs. nonpresented), relatedness of the study list items to the critical word (related vs. unrelated), and recall phase (precritical vs. postcritical). In addition, a fourth within-subjects variable was included: input position (5th, 11th, and 17th).

**Participants**

Forty-three participants were tested in this experiment. They were all student volunteers who fulfilled a requirement for their introductory psychology course by participating in the experiment.

**Materials and procedure**

All of the procedural details were the same as in the previous experiments. With regard to materials, this experiment was most like Experiment 2. The lists used were once again a mixture of related and unrelated words and were 21 words long if in the presented...
condition and 20 words long if in the nonpresented condition. There were no buffer items. However, twelve DRM lists were used instead of eight. The additional lists were the needle, anger, trash, and soft lists. Each participant saw six lists in each presentation condition (presented vs. nonpresented). Across the six lists in the presentation condition, the presented critical words were presented at each of the input positions (i.e., 5th, 11th, and 17th) in two of the lists. For the most part, the same unrelated words were used in eight of the lists as had been used in Experiment 2, with some minor changes. For example, in one list in Experiment 2, both blue and pink were unrelated to the gist of the list, but inasmuch as they are both color words, they are categorical associates. “Blue” was replaced by “box” in this experiment.

Results and discussion

Of the 43 participants tested, 39 (91%) accurately recalled a presented critical word at least once and falsely recalled a nonpresented critical word at least once. Of these 39, 36 recalled at least one presented critical word that had been presented at each of the three input positions. Across the 36 participants, there were 216 opportunities to recall presented critical words and 216 opportunities to recall nonpresented critical words. Participants recalled presented critical words 185 times, or an average of 5.1 times out of every 6 lists (86%). In the presented condition, average recall of related words was 50% and of unrelated words was 26%. Participants recalled nonpresented critical words 98 times, or an average of 2.7 times out of every 6 lists (45%). In the nonpresented condition, average recall of related words was 52% and of unrelated words was 27%. In short, overall performance in Experiment 5 was quite similar to that in Experiment 2.

Critical word output position as a function of presentation condition and mean recall as a function of relatedness, recall phase, and presentation is displayed in Table 1. The average absolute output position of presented critical words was 4.3 and the average absolute output position of nonpresented critical words was 5.8. This difference was significant, \( t(35) = -3.58, SEM = .407 \); 95% CI = 6.2–2.3. As in the previous five experiments, precritical recall was less, and postcritical recall greater, in the presented condition than in the nonpresented condition, producing a significant 2-way interaction of presentation and recall phase, \( F(1,35) = 14.17, MSE = 2.171 \). The difference between total recall in the presented and nonpresented conditions (8.6 vs. 8.9) was not significant (\( t < 1.1 \)). When the difference between relative output position in the presented and nonpresented conditions (.51 vs. .63) was examined, this was also significant, \( t(35) = -2.86, SEM = .043 \). In short, there was strong support for the output position hypothesis.

Most important for present purposes was whether output position of a presented critical word varied as a function of input position. Whether the critical word was presented in the 5th, 11th, or 17th position in a 21-item list had little effect on its output position. The means were 4.2, 4.5, and 4.3, respectively, \( F < 1 \).

To analyze the power of this experiment, the following rationale was used. This experiment investigated the possibility that the central output position for presented critical words was due to their central input position. More generally, this is a claim that output position recapitulates input position. This claim predicts that an early input position should result in an early output position and a late input position should result in a late output position. How early and how late? The 5th input position in the 21-item study list used in this experiment is 24% through the study list; the 17th input position is 81% through the study list. Mean total recall in the presented condition was 8.6 items. Thus, a strong correspondence between input and output positions would be reflected in an output position for the early input item of 24% of 8.6, or 2.05, and an output position for the late input item of 81% of 8.6, or 6.95. The mean output position for the 11th (central) input position was 4.5. Thus, the differences between the (hypothesized) early and late output positions and the (actually obtained) central output position were both 2.45. The obtained standard deviation for the differences between the output positions for the 5th and 11th items in this experiment was 2.45; for the 11th and 17th items, it was 3.06. These yielded Cohen's \( d \)'s of 1.00 and .80, respectively. The power of this experiment to detect effects of this size exceeded .99 (\( N = 36, \alpha = .05 \), directional test). The power to detect effects just half these sizes (corresponding to hypothesized output positions of 3.28 and 5.73 and a mean difference from the central output position of 1.23) was .91 and .77, respectively. In short, it seems that the absence of an effect of a presented critical word’s input position on its output position was not due to insufficient power. Instead, it appears that the relatively central output position of the presented critical word in these experiments was due to the notion of cognitive triage during the initial direct access phase in a recall protocol.

As in Experiment 4, the relatedness hypothesis was again supported. Collapsing across presentation, the advantage for related recall over unrelated recall in the postcritical phase (1.5) was greater than that same advantage in the precritical phase (1.0), \( F(1,35) = 4.78, MSE = .842 \). The difference between recall of related (2.7) and unrelated (1.2) words during the reconstructive phase was significant, \( F(1,35) = 99.11, MSE = .784 \). The difference between recall of related (2.5) and unrelated (1.2) words during the reconstructive phase was significant, \( F(1,35) = 55.95, MSE = .639 \). Compared with Experiment 2, the principal change in the data
pattern that resulted in a significant relatedness by recall phase interaction in this experiment was the much smaller difference between precritical related and unrelated recall in this experiment, especially in the presented condition.

The 3-way interaction was significant, $F(1,35) = 5.60$, $MSE = .768$. Again, the data pattern appeared to be an artifact of not counting recall of the critical word as part of either the direct access or reconstructive phase. The 2-way interaction of presented and relatedness was not significant, $F < 1$. The main effect of relatedness was once again significant, $F(1,57) = 188.30$, $MSE = .581$. Participants recalled almost twice as many related words (5.1) as unrelated words (2.7). Finally, participants recalled an equal number of words before and after the critical word (both $M's = 3.9$).

**General discussion**

There have been a number of reports noting that nonpresented critical words are output relatively late during free recall of DRM lists (McDermott, 1996; Payne et al., 1996; Roediger & McDermott, 1995). Both dual-retrieval theory and activation/monitoring theory can account for this observation.

In the dual-retrieval approach (Brainerd et al., 2002), information is encoded in both a verbatim format and a gist format. In a free recall test, veridical memories can be produced via direct access of a verbatim trace or via reconstruction from a gist trace, but false memories are produced only via the latter. The dual-retrieval approach argues that direct access constitutes the initial phase during free recall and that stronger traces are output last during this phase (i.e., the notion of cognitive triage). Reconstructive retrieval constitutes the subsequent phase during free recall. Stronger traces are output first during this phase. From the perspective of this theory, the relatively late output of nonpresented critical words is consistent with the notion that reconstructive retrieval is the later of the two retrieval processes.

In the activation/monitoring approach, the recall of nonpresented critical words is a function of their activation at study and the failure of monitoring mechanisms at test. From the perspective of this theory, the relatively late output of nonpresented critical words occurs because their activation is not as strong as it is for words that were actually presented, which are output in reverse order of their activation strength, with the most activated being output first.

However, a direct comparison of nonpresented critical words with presented critical words is relatively rare (but see McDermott, 1997; Miller & Wohlford, 1999). Both theories would predict that the output position of presented critical words should be prior to that for nonpresented critical words. In dual-retrieval theory, presented critical words are more likely than nonpresented critical words to be output during the direct access phase, which occurs prior to the reconstructive phase. In activation/monitoring theory, presented critical words yield more strongly activated traces than do nonpresented critical words and will be output prior to nonpresented critical words if the most activated traces are output first.

However, these theories diverge in their predictions about when presented critical words should be output. Because it incorporates the notion of cognitive triage, dual-retrieval theory predicts that presented critical words should be output last during the direct access phase. As a result, the output position for presented critical words should not only be both prior to that for nonpresented critical words, but it should occupy a relatively central output order position. In contrast, because it incorporates the notion that output order is an inverse function of activation strength, activation/monitoring theory predicts that presented critical words should not only be prior to that for nonpresented critical words, but should be output first during recall.

Finally, the two theories also differ in the precision with which they predict the output position of nonpresented critical words and, by extension, the relationship between the output positions of presented and nonpresented critical words. In dual-retrieval theory, the direct access phase and reconstructive phase occur in succession. Presented critical words should be output last during the direct access phase. Nonpresented critical words should be output first during the reconstructive phase. As a result, the output positions of presented and nonpresented critical words should differ by about one word or one position. In contrast, although a fairly precise prediction about the output position of presented critical words—that they should be output first in recall—can be derived from activation/monitoring theory, a comparably precise prediction about the output position of nonpresented critical words is unavailable, other than to say that it should occur relatively late in recall.

Collapsing across the six experiments reported here, the average output position of presented critical words was 3.9 out of 8.4 words recalled and the average output position of nonpresented critical words was 4.9 out of 8.2 words recalled. Dual-retrieval theory can account for many different aspects of this data pattern. That is, it can account for the relatively late output of nonpresented critical words, the prior output of presented critical words, the relatively central output position of the presented critical words, and the one word/position separation between the output positions for presented and nonpresented critical words. In contrast, although activation/monitoring theory can account for the relatively late output of nonpresented critical words and the prior output of presented critical words, it cannot account for the relatively central output position of...
presented critical words or the one word/position separation of the output positions for those two types of words.

One alternative explanation for the relatively central output position of presented critical items in Experiments 1–3 attributed it to the early output of the primacy and recency portions of the study lists, which would relegate output of presented critical words to a more central output position. In Experiment 4, primacy and recency buffers were included in the lists and then left unscored. Nevertheless, the output position of presented critical items remained much closer to the center, than to the beginning, of the recall protocol.

Another alternative explanation for the relatively central output position of presented critical words in Experiments 1–4 attributed it to the fact that the input position of presented critical words was always central in those experiments. In Experiment 5, the input position of presented critical words was varied. Nevertheless, the output position of presented critical words did not significantly vary as a function of input position, remaining relatively central. At first glance, this finding appears to contradict the notion that recall relies on seriation or input–output correspondence (e.g., Nairne et al., 1991). However, as Nairne et al. (1991) recognized, the reliance of recall on seriation decreases when list items are categorically related (e.g., Murdock & Vom Saal, 1967). In the experiments reported here, list items were assortatively, rather than categorically, related. However, one possibility is that this aspect of the study lists impaired subjects’ typical reliance on order information, which usually yields a correspondence between input and output order and, as a result, subjects relied on an alternative mechanism—that is, cognitive triage—for controlling their output during recall.

In short, the output position hypothesis derived from dual-retrieval process theory was supported in all five experiments reported here. In contrast, it appears that the most straightforward extension of activation/monitoring theory is inconsistent with the finding that presented critical items are not output earlier. In that extension, output order is inversely related to trace strength, which predicts that the presented critical item should be output first.

It is interesting to consider whether a combination of activation/monitoring theory and cognitive triage could account for the output position findings. Once it is accepted that presented noncritical items are output first so as to avoid output interference, then the difference between the output positions of presented critical words and nonpresented critical words is easily attributable to a difference in the degree to which they were activated at study. That is, nonpresented critical words generate implicit associative responses (IAR’s), but the trace strength of these items is not as great as that for presented critical items. In this account, the output of the last trace during the first phase signals the need for a shift in response criterion, not a shift in retrieval process. The nonpresented critical item is output first during the second phase because, of all those traces that were not strong enough to be output during the first phase, it is the most likely to slip through a monitoring process because of: (a) its similarity to presented noncritical items, a similarity that was generated when the IAR was produced at study, and (b) because of its trace strength. Note that, in this approach, the notion of reconstructive processing from gist memory at test—an important component of the dual-retrieval theory—is not necessary to account for the relationship between the output order of presented and nonpresented critical items. Finally, such an account can also explain the negative correlation between trace strength and output order during the second phase of recall by attributing it to an increasingly liberal criteria for response. Thus, it appears that activation/monitoring theory, with some modifications, could account for the output position data reported here.

A second hypothesis derived from dual-retrieval theory—the relatedness hypothesis—was also investigated in Experiments 2–5. Along with words related to the gist of the list, words unrelated to the gist of the list were included in the study lists. The dual-retrieval approach argues that there are two retrieval processes—direct access of verbatim traces and reconstruction from gist traces—that occur consecutively during a recall test. During reconstructive retrieval, words related to the gist of the list should be generated much more often than words unrelated to the gist and, as a result, the number of related words that exceed the judgment criterion for response should far outstrip the number of unrelated words. In contrast, verbatim traces primarily entail perceptual information and successful direct access of these traces is a function of the extent to which that information has been successfully encoded. As such, during direct access, the advantage for recall of words related to the gist of the list should be considerably less than in the reconstructive phase. In short, the relatedness hypothesis stated that the advantage for related word recall should be greater after a critical word is output than before.

As noted in the introduction, a quite different pattern of results would be predicted by activation/monitoring theory. Assuming that presented/related words would produce more highly activated memory traces than presented/unrelated words, and that the most highly activated traces would be output first, activation/monitoring theory would predict that there should be relatively large advantage for related word recall early in the output and that this advantage should decrease, rather than increase, as recall progressed.

Support for the relatedness hypothesis was less consistently observed than was support for the output position hypothesis. The advantage for related word
recall in the postcritical phase was larger than in the pre-critical phase in Experiments 4 and 5. In Experiments 2 and 3, the advantage for related word recall was equivalently large in both the precritical and postcritical phases. In Experiment 2, there was a nonsignificant trend in the direction predicted by dual-retrieval theory. In Experiment 3, there was a nonsignificant trend in the direction predicted by activation/monitoring theory. In sum, across the four experiments in which the relatedness hypothesis was tested, there was more support for the dual-retrieval prediction than for the activation/monitoring prediction.

There did not appear to be any systematic difference between those experiments in which the relatedness effect was observed and those in which it was not. The effect was observed in Experiment 4, after having not been observed in Experiments 2 and 3. There were two major differences between the latter experiment and the pair of earlier experiments. First, in Experiment 4, the unrelated words were categorically related across the lists, whereas in Experiments 2 and 3 they were either unrelated or associatively related across the lists, respectively. Second, buffer items were included in Experiment 4, but not in Experiments 2 and 3. But neither of these elements (i.e., unrelated words that were categorically associated across the lists or buffer items) was present in Experiment 5, where the effect was once again observed. The only differences between Experiment 5 and Experiment 2 were the addition of two lists and some minor tweaking of the unrelated words. The vast majority of the unrelated words remained the same across those two experiments.

Another possibility was that the difference between related and unrelated word recall was proportional to the total number of words recalled in that phase. For example, in Experiment 4, where the postcritical difference between related and unrelated word recall was greater than the precritical difference, postcritical recall was greater than precritical recall. However, this was also true in Experiments 2 and 3, where the relatedness effect was not observed. In addition, recall was equal across the precritical and postcritical recall phases in Experiment 5, but the relatedness effect was again observed. In short, the relatedness effect did not appear to be dependent on the relative amount of total recall in the two phases.

Another aspect of the relatedness effect that merits discussion was the observation that, across the four experiments in which the relatedness hypothesis was tested, the average related word advantage before output of a critical word was 1.0. (The average related word advantage after output of the critical word was 1.4.) That is, it is clear that there was a sizable difference between recall of related and unrelated words prior to critical word output. In some ways, this observation seems a bit at odds with the idea that, prior to this point in a recall protocol, recall is driven by direct access to verbatim traces. Theoretically, one of the most important aspects of verbatim traces is that they consist predominantly of surface (i.e., perceptual) information. For this reason, it is not immediately apparent why information regarding meaning (i.e., relation to the gist of the list) should yield better performance during this recall phase.

One possible explanation has to do with process-pure assumptions. In its original formulation, Jacoby (1991) described the process-pure assumption as one in which researchers assumed that particular memory tests were sensitive to only one retrieval process. He argued that such assumptions were probably incorrect and introduced the process dissociation procedure as a method for separating the contributions of two processes—familiarity and recollection—to recognition test performance. Although this particular method for dissociating the contribution of multiple processes to test performance has been criticized and alternative methods have been proposed (e.g., Brainerd, Reyna, & Mojardin, 1999), the basic idea that memory tests are often not process pure has been fairly well accepted. The same might be true for the two phases of recall in the dual-retrieval process theory. For explanation purposes, dual-retrieval theory may be described such that the only retrieval process occurring before (and during) recall of a critical presented word would be direct access and the only retrieval process occurring (during and after recall of a critical nonpresented word would be reconstruction. Another possibility is that some reconstructive retrieval also occurs during the first phase and that some direct access retrieval also occurs during the second phase. This would account for the advantage in related word recall during the first phase and for the presence of recall of unrelated words during the second phase.

Another possibility is that the process-pure assumption is a correct assumption when applied to dual-retrieval theory and that the notion of verbatim traces must be re-conceptualized. One of the assumptions underlying the relatedness hypothesis was that relational processing would have relatively little bearing on the strength of verbatim traces. Perhaps this assumption was overstated. There are two lines of argument on this point. First, even though verbatim traces are typically conceptualized as being more closely associated with item information, perhaps this does not necessarily exclude the possibility that their strength is somewhat a function of relational processing. Indeed, Einstein and Hunt (1980) have argued that individual item information can be influenced by relational processing. Second, in dual-retrieval theory (e.g., Brainerd et al., 2002), one of the most important aspects of verbatim traces is that these traces consist predominantly of surface (i.e., for visual word stimuli, letter and word forms)
information. However, relational processing can promote elaborative processing (i.e., the process of relating semantic information from the stimulus event to semantic memory; Craik & Tulving, 1975) and it has been argued that elaborative processing can promote the integration of surface information (e.g., Mandler, Hamson, & Dorfman, 1990). In short, there are at least a couple of reasons to think that relational processing at study could increase the strength of verbatim traces and, as in the experiments here, lead to an advantage for related recall even during the direct access phase.

In summary, two hypotheses—the output position hypothesis and the relatedness hypothesis—were investigated. With regard to the output position hypothesis, the results were very consistent across the experiments. Nonpresented critical words were output later than presented critical words, but both were output at relatively central positions during recall. This pattern of results was accounted for in a straightforward manner by dual-retrieval theory. In this theory, there are two successive phases during recall: direct access of verbatim traces and reconstructive retrieval from the gist of the list. These findings also corroborated the theory’s prediction of a one word/position difference between the output position of presented and nonpresented critical words. In general, this pattern of results was inconsistent with a straightforward extension of activation/monitoring theory, but could perhaps be accounted for by a version of activation/monitoring theory that incorporated the notion of cognitive triage, i.e., strongest traces are output last, amongst other modifications.

With regard to the relatedness hypothesis, the results were much less consistent. In two experiments, the difference between related and unrelated word recall was significantly greater after the output of critical items than before the output of critical items, but not in two other experiments. In short, the effect is probably real, but small. This was unfortunate because this hypothesis also differentiated dual-retrieval and activation/monitoring theories. Principally, this was because of the incompatibility between the emphasis that activation monitoring theory places on constructive processing during study and the emphasis that dual-retrieval theory places on reconstructive processing at test. That is, of the two approaches, only dual-retrieval theory—because it makes the claim that the second of the dual-retrieval processes is reconstructive processing from gist memory—would make the prediction that the difference between postcritical recall of related and unrelated words should be greater than the difference between precritical recall of related and unrelated words.

One of the themes sounded in the introduction had to do with the similarity of veridical and false memories. It was pointed out that the two theories differed on this count, with activation/monitoring theory tending toward the view that isomorphic memory traces give rise to veridical and false memories in the DRM paradigm. In contrast, dual-retrieval theory offers a clear distinction between the memory traces retrieved during the direct access phase and the memory traces that give rise to reconstructive retrieval and, at times, false recall. The results of the present experiments can be added to the list of ways in which veridical and false memories differ: False memories for nonpresented words are output later than veridical memories for those same words when they are presented. Consequently, to the extent that activation/monitoring theory tends toward the idea that veridical and false memories are isomorphic, this finding is more consistent with dual-retrieval theory. To be clear, there is no evidence in the present experiments that directly speaks to the difference between traces upon which veridical and false memories depend. However, the one-word difference in output position for veridical and false memories of critical words and the relatedness effect were predictions drawn from dual-retrieval theory which, in turn, was predicated on the notion that different types of traces (i.e., verbatim and gist) are retrieved during successive phases (i.e., direct access and reconstruction) in recall. Thus, from this perspective as well, the present results are more consistent with the notion that the memory traces subserving veridical and false memories are not isomorphic.

In conclusion, several observations regarding output order during free recall converged to support a dual-retrieval theory of free recall (Brainerd et al., 2002). Some of these findings can be explained by a modified activation/monitoring approach that incorporated the notion of cognitive triage. However, dual-retrieval theory also appeared better equipped to account for the finding—to the extent that it is reliable—that the difference between recall of related and unrelated words was greater after output of the critical item (i.e., during the reconstructive phase), than before output of the critical item (i.e., during the direct access phase). This is not to say that there are not other mechanisms that could possibly account for the relatedness effect. For example, it is possible that as recall progresses, recall of related words, but not unrelated words, accelerates because of “priming” from prior recall of related words. Further research is needed to establish the reliability of, and reasons for, the relatedness effect.

References


