Falsely recalled items are rich in item-specific information

DANIEL J. BURNS, CARIN L. JENKINS, AND ERICA E. DEAN
Union College, Schenectady, New York

Current theorizing suggests that critical lures in the Deese/Roediger–McDermott (DRM) procedure are often falsely remembered because they have received considerable relational processing (e.g., spreading activation or encoding of gist information). We used a repeated-testing paradigm to assess the amount of item-specific and relational processing given to the list items and the critical lures. Research has shown that items receiving item-specific processing are more likely to be recovered across successive tests. They are also output more slowly but more steadily throughout the recall period. In two experiments, we manipulated the processing performed on list items and then used item gains and cumulative recall curves to assess the amount of item-specific and relational information encoded for both list items and lures. The results suggest that increasing the relational processing of list items increased item-specific processing of lures, whereas increasing item-specific processing of list items decreased item-specific processing of lures. We conclude that critical lures are typically rich in item-specific information, relative to list items.

Falsely remembering an event that never occurred but that may be related to events that did occur is termed illusory recollection. The Deese/Roediger–McDermott (DRM) procedure has been used extensively to study factors that influence illusory recollection (Deese, 1959; Roediger & McDermott, 1995). In the DRM procedure, participants are presented with a list of items (e.g., "ice" and "wet") that are all associatively related to a nonpresented critical lure (e.g., "cold"), followed by a test of recall or recognition of the list items. Not only is the critical lure often falsely remembered, but participants may be highly confident in the veracity of the false memory (see, e.g., Read, 1996) and may falsely remember specific details about the item’s presentation, including its list position and the modality in which it was presented (see, e.g., Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997; Payne, Elie, Blackwell, & Neuschatz, 1996).

Although there are few boundary conditions for the false memory effect in the DRM procedure, several variables have been shown to reduce false recollection, including the use of pictures rather than words as list items and visual rather than auditory presentation (see Roediger, McDermott, & Robinson, 1998). Several theories have been fashioned to explain these reductions in false recollection. The leading theories may differ in detail, but all assume that item-specific processing of the list items is a major cause. Moreover, these theories all assume that the additional item-specific information is used, in one way or another, to distinguish list items from critical lures. For example, the activation/monitoring framework assumes that the encoding of item-specific features of the list items allows for more accurate source monitoring (see, e.g., McDermott & Watson, 2001; Roediger, Watson, McDermott, & Gallo, 2001). Fuzzy-trace theory suggests that item-specific processing of the list items typically results in less reliance on gist traces during retrieval, thereby reducing false recollection (see, e.g., Brainerd & Reyna, 1993, 1998; Reyna & Brainerd, 1995). Proponents of the distinctiveness heuristic suggest that when sufficient item-specific features of the list items are encoded, participants call upon a stricter decision criterion, one that demands access to the item-specific features (see, e.g., Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999).

Extant theories also suggest that the critical lures receive a considerable amount of semantic, associative, or relational processing. For example, spreading-activation approaches state that critical lures receive semantic activation from the related list items. Other theories assume that a considerable amount of gist information about the critical lures is encoded (see, e.g., Brainerd & Reyna, 1993, 1998). What current explanations have not considered in detail, however, is precisely how the list items are processed differently from the critical lures. It may seem obvious that the list items receive item-specific processing by virtue of their physical presentation, whereas critical lures receive primarily relational processing (or activation) via spreading activation from the related list items. Although this is both a highly intuitive and generally accepted (albeit largely untested) view, recent evidence supports the opposite conclusion. Specifically, Burns, Martens, Bertoni, Sweeney, and Lividini (2006) presented evidence that critical lures actually received more item-specific

D. J. Burns, burnsd@union.edu

Copyright 2007 Psychonomic Society, Inc.
processing than did list items, whereas list items received more relational processing than critical lures did.

Burns et al. (2006) used several indices of item-specific and relational processing to assess encoding differences between list items and lures. Following the presentation of a series of DRM lists, three successive 4-min recall tests were given. This repeated-testing procedure allowed the researchers to measure both item gains (the recall of previously unrecovered items) and item losses (the failure to recall previously recalled items). It is now well documented that increases in item gains are associated with increases in item-specific processing, whereas decreases in item losses across successive tests are associated with increases in relational processing (see, e.g., Burns & Gold, 1999; Burns & Hebert, 2005; Engelkamp & Seiler, 2003; Klein, Loftus, Kihlstrom, & Aserson, 1989; Klein, Loftus, & Schell, 1994; Mulligan, 2000, 2001, 2002; Mulligan & Duke, 2002).

A decrease in item losses results from enhanced relational processing, apparently because the encoded relational information provides an organized search strategy that facilitates accessibility of the items (see Burns & Gold, 1999; Burns & Schoff, 1998; Hunt & McDaniel, 1993; McDaniel, Moore, & Whitman, 1998). Presumably, the same relational cues are used across successive tests, resulting in a similar retrieval strategy that minimizes item losses. Furthermore, the likelihood that participants will use the same retrieval routes across tests increases on later tests (McDaniel et al., 1998), typically resulting in progressively fewer losses across successive tests (see, e.g., McDaniel et al., 1998; Mulligan, 2001, 2002; Mulligan & Duke, 2002).

The theoretical connection between item-specific processing and item gains is that item-specific processing produces numerous potential retrieval cues for each item. Since it is unlikely that all of the cues will be exhausted by the end of the first recall period, many of these cues may be accessed on later tests, producing previously unrecovered items (Klein et al., 1989). Alternatively, it may be that following item-specific processing, the specific items that become accessible on any particular test are much more variable because recall is not guided by an organized retrieval strategy. Thus, items not accessed on the first test may be accessed on the second test and vice versa (see, e.g., Hunt & McDaniel, 1993).

In addition to recording item gains and losses, Burns et al. (2006) plotted the cumulative percentage of items recalled across all 12 min of the three 4-min tests. It turns out that relational and item-specific processing produce very different cumulative recall curves. These cumulative recall curves are well described by the exponential equation

\[ n(t) = n(\infty) \left(1 - e^{-\lambda t}\right), \]

where \( n(t) \) is the number of items recalled at time \( t \), \( n(\infty) \) is asymptotic recall level, \( e \) is the base of the natural logarithm, and \( \lambda \) is the rate of approaching asymptote (see, e.g., Bousfield & Sedgewick, 1944; Roediger, Stellon, & Tulving, 1977). Asymptotic recall level and rate of approaching asymptote have consistently been found to be inversely related, such that higher levels of asymptotic recall produce a slower approach to asymptote (see, e.g., Bousfield & Sedgewick, 1944; Herrmann & Chaffin, 1976; Herrmann & Murray, 1979; Johnson, Johnson, & Mark, 1951; Kaplan, Carvellas, & Metlay, 1969). However, Burns and Schoff (1998) showed that this relationship is often violated when differing levels of item-specific and relational processing are performed. The curves typically produced following item-specific, relational, or both types of processing have been published in several places (Burns, 2006; Burns & Hebert, 2005; Burns et al., 2006; Burns & Schoff, 1998), so we will not reproduce them here. Suffice it to say that relational processing, presumably owing to the use of an organized retrieval strategy, results in initially rapid recall that reaches asymptote very quickly, whereas item-specific processing produces a slower but steadier rate of recall.

Burns et al. (2006) showed that critical lures produced more item gains than did list items. They also demonstrated that the cumulative recall curves produced for the critical lures were slower but steadier than those produced for list items. However, only the list items showed a decrease in item losses across successive tests. These results suggest that the critical lures actually received more item-specific processing but less relational processing than did the list items, a conclusion that is clearly contrary to current theorizing.

How is it possible that nonpresented critical lures receive more item-specific processing than do list items that are physically presented? Burns et al. (2006) suggested two-step priming as one possible mechanism (e.g., see Balota & Lorch, 1986). It is possible that the list items activate the critical lure, which in turn activates concepts related to the critical lure but unrelated to the list items. For example, the list words “ice” and “wet” might spread activation to the critical lure “cold,” which then activates the related concept “beer.” Although related to the lure, “beer” is not related to the list items. Thus, “beer” would be an item-specific retrieval cue that could facilitate recall of the critical lure.

If critical lures receive more item-specific processing than do list items, then current theories are at a loss to explain why increasing item-specific processing of the list items decreases false recollection (see, e.g., Arndt & Reder, 2003; Smith & Hunt, 1998). These theories assume that the extra item-specific information given to the list items serves as the basis for discrimination at time of test. If the critical lures are also rich in encoded, item-specific information, then there would be no basis for discriminating between the list items and the lures. Other theories, referred to as the impoverished relational encoding explanations (see Hege & Dodson, 2004; Schacter, Cendan, Dodson, & Clifford, 2001), assume that increasing item-specific processing of the list items results in a decrease in relational processing given to the critical lures, thereby reducing false recollection. However, if, as Burns et al. (2006) assume, the critical lures typically receive little relational processing, then there would be very little to impoverish.

Here we present two experiments that further explored the type of information typically encoded for list items.
and for critical lures. Burns et al. (2006) assumed that the item-specific processing given to the critical lures was the result of spreading activation from the list items. This leads to the prediction that manipulations that increase the amount of spreading activation should enhance the item-specific processing given to the lures. To test this prediction, in Experiment 1, we varied the associative strength between the list items and the critical lures, assuming that an increase in associative strength would increase activation of the critical lures (see, e.g., Hancock, Hicks, Marsh, & Ritschel, 2003).

In Experiment 2, we directly tested the contention of Burns et al. (2006) that list items receive primarily relational processing. A prediction derived from this view is that encoding tasks designed to promote relational processing of the list items would lead to processing that is redundant with the processing typically performed (i.e., relational processing) and would therefore have little effect on recall of either the list items or the critical lures. However, encoding tasks designed to promote item-specific processing of the list items would lead to processing that is different from the relational processing that is typically performed, thus affecting recall of both the list items and the lures. We tested this prediction by requiring some participants to perform relational or item-specific orienting tasks on the DRM list items during presentation.

**EXPERIMENT 1**

Experiment 1 manipulated the associative strength between the list items and the critical lures. We intermixed lists of items that were either strongly or weakly associated with the critical lures. Previous research has shown that strongly associated items produce more false memories (see, e.g., Gallo & Roediger, 2002; McEvoy, Nelson, & Komatsu, 1999). Here, we used the repeated-testing paradigm and a cumulative recall analysis to assess the effect of increasing associative strength on the amount of item-specific and relational processing given to the list items and the lures. We assumed that strong associates would result in greater activation of the critical lures (see, e.g., Hancock et al., 2003; Tse & Neely, 2005). Of import was the effect that increasing associative strength would have on the participants’ responses to the critical lures in terms of item gains, item losses, and cumulative recall. If, as current theory suggests, increasing activation increases relational processing (between the list items and each lure), then it would result in a decrease in item losses across successive tests and a rapid approach to asymptotic recall levels. If, however, increasing associative strength increases item-specific processing of the critical lures, then it would produce an increase in item gains and recall curves that approach asymptotic recall levels more slowly.

**Method**

**Participants and Materials.** The 20 undergraduate students from Union College who participated in Experiment 1 were given credit toward an out-of-class activity requirement of their introductory psychology course. They were tested in groups of 5 or fewer. All participants were presented a list of 100 words composed of 10 sublists. The 10 sublists were taken from the lists used by Roediger and McDermott (1995). The items selected for half of the sublists were the strongest 10 associates, and for the other half they were the 6th–15th strongest associates. Counterbalancing across participants ensured that the 5 DRM lists used as strong associate sublists for half of the participants were weak associate sublists for the remaining half of the participants. The associates from each sublist were blocked during presentation, and the items within each sublist were ordered from strongest to weakest associative strength.

**Procedure.** Students were informed that they would hear a list of words being presented at a rapid rate and that they would be asked to recall the words afterward. No mention was made of the repeated tests. A tape recording was played, and participants heard the 100 items, which were presented at a rate of 1 every 1.5 sec. The recall instructions following list presentation were read aloud by the experimenter and lasted about 60 sec. The students were told to write as many of the words that they had heard as they could remember, in any order they liked. They were instructed to be reasonably certain each word had been presented before writing it down. At the end of each minute of recall, participants were asked to draw a line under the last word that they had written down, which allowed cumulative recall performance to be tracked.

The first recall test lasted 4 min. The recall sheets were then collected, and students were given a 4-min distractor task (requiring written recall of the names of the U.S. states), followed by the second 4-min recall test, which was identical to the first. A second 4-min distractor task (writing down the names of state capitals) was given before the third and final recall test, which was identical to the first two tests.

**Results**

To facilitate comparisons between the list items and the critical lures, all raw scores were converted to percentages. The various performance measures for both the list items and the critical lures are presented in Table 1, which shows that both the list items and the critical lures from the strong associate lists were recalled better than those from the weak associate lists. Separate 2 (associative strength) × 3 (test number) within-subjects ANOVAs were conducted on the list items and the critical lures. For the list items, only the main effect of associative strength was significant [$F(1,19) = 15.11, M_\text{S}\text{.e.} = 159.68$]. Neither the main effect of test number nor the interaction approached significance [$F(2,38) = 2.11, M_\text{S}\text{.e.} = 7.07$, and $F(2,38) = 1.47, M_\text{S}\text{.e.} = 8.79$]. For the critical lures, the associative strength × test number interaction reached significance [$F(2,38) = 3.27, M_\text{S}\text{.e.} = 92.81$]. Follow-up one-way ANOVAs showed that whereas critical lure recall increased across successive tests for the strong associate lists [$F(2,38) = 3.41, M_\text{S}\text{.e.} = 148.72$], there was not a corresponding increase for the weak associate lists ($F < 1$).

The percentage of items gained was calculated by dividing the raw number of items gained from one test to the next by the total number of items presented and then multiplying the quotient by 100. We also divided the number of item gains by the number of items not recalled on the previous test. The second procedure accounts for the fact that conditions with higher initial recall levels have fewer possible item gains on later tests. Except where noted, the two item gain formulas produced an identical pattern of significant results, so we report only the results from the first formula. On the relatively rare occasions on which an item was recalled on the first test, forgotten on
Table 1

Mean Performance Measures (in Percentages) and Standard Deviations for the List Items and Critical Lures in Experiment 1 As a Function of Associative Strength

<table>
<thead>
<tr>
<th>Type of Item</th>
<th>Weak Associates</th>
<th></th>
<th></th>
<th></th>
<th>Strong Associates</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List</td>
<td>Critical</td>
<td>List</td>
<td>Critical</td>
<td>List</td>
<td>Critical</td>
<td>List</td>
<td>Critical</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Test 1 recall</td>
<td>19.70</td>
<td>10.86</td>
<td>10.00</td>
<td>16.54</td>
<td>27.40</td>
<td>11.90</td>
<td>25.00</td>
<td>20.39</td>
</tr>
<tr>
<td>Test 2 recall</td>
<td>18.10</td>
<td>10.08</td>
<td>9.00</td>
<td>13.73</td>
<td>27.40</td>
<td>10.24</td>
<td>29.00</td>
<td>25.53</td>
</tr>
<tr>
<td>Test 3 recall</td>
<td>19.00</td>
<td>11.40</td>
<td>9.00</td>
<td>13.73</td>
<td>28.90</td>
<td>10.73</td>
<td>35.00</td>
<td>22.36</td>
</tr>
<tr>
<td>Test 2 gains</td>
<td>1.80</td>
<td>2.59</td>
<td>2.00</td>
<td>6.16</td>
<td>3.10</td>
<td>2.63</td>
<td>8.00</td>
<td>11.96</td>
</tr>
<tr>
<td>Test 3 gains</td>
<td>1.50</td>
<td>1.70</td>
<td>0.00</td>
<td>0.00</td>
<td>1.80</td>
<td>2.42</td>
<td>7.00</td>
<td>11.74</td>
</tr>
<tr>
<td>Test 2 losses</td>
<td>3.40</td>
<td>4.01</td>
<td>3.00</td>
<td>9.79</td>
<td>3.10</td>
<td>3.28</td>
<td>4.00</td>
<td>10.46</td>
</tr>
<tr>
<td>Test 3 losses</td>
<td>0.80</td>
<td>1.36</td>
<td>1.00</td>
<td>4.47</td>
<td>1.80</td>
<td>2.50</td>
<td>4.00</td>
<td>10.46</td>
</tr>
<tr>
<td>Approach to asymptote (λ)</td>
<td>0.73</td>
<td>0.74</td>
<td>0.74</td>
<td>0.83</td>
<td>0.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymptote [n(∞)]</td>
<td>21.59</td>
<td>11.35</td>
<td>29.95</td>
<td>36.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note—Test 3 recall does not always equal Test 2 recall plus Test 3 gains minus Test 3 losses, because items lost on the second test but gained on the third test were not included as item gains. Approach to asymptote refers to the estimated rate of approach to the asymptotic cumulative recall level using Equation 1 and is based on group means.

The second test, and then recalled again on the third test, it was not considered an item gain. Separate 2 (associative strength) × 2 (test number: Test 2 vs. Test 3) ANOVAs were conducted on the list item and critical lure gains. The ANOVA for the list items produced no significant effects (largest $F = 2.67$, for the main effect of test number). For the critical lures, however, the strong associates produced more item gains than did the weak associates [$F(1,19) = 12.79$, $MS_e = 66.05$]. Neither the main effect of test number nor the interaction approached significance (both $Fs < 1$).

The item loss percentages were also calculated by dividing the raw number of items lost from one test to the next by the total number of items presented and then multiplying the quotient by 100. As can be seen in Table 1, item losses were rare, with no condition losing more than 4% of items on any test. Presumably because of this floor effect, the 2 (associative strength) × 2 (test number: Test 2 vs. Test 3) ANOVA conducted on the list item and critical lure losses failed to produce any significant effects (largest $F = 1.13$).

Cumulative recall curves for all 12 min of recall (across the three tests) for both the list items and the critical lures as a function of associative strength are shown in Figure 1. Whereas the other three conditions approached asymptote fairly quickly, the critical lures from the strongly associated lists produced a slower and steadier approach to asymptote, suggesting that they received primarily item-specific processing. Using the cumulative recall data shown in Figure 1, we estimated $\lambda$ and $n(\infty)$ for each of the two types of items.2 These estimates, presented in Table 1, show that the critical lures from the strong associate lists produced a much lower estimate of $\lambda$ than did the other three groups.

Discussion

Replicating the work of previous researchers (e.g., McEvoy et al., 1999), we found that strongly associated lists produced higher recall of both the list items and the critical lures. These results indicate that our manipulation of associative strength was effective. The recall percentages also demonstrate that only recall of the critical lures from the strongly associated lists increased across tests (hypermnesia). Burns et al. (2006) used strongly associated lists and also demonstrated hypermnesia for the critical lures but not for the list items (see also McDermedott, 1996; Thapar & McDermedott, 2001; and Toglia, Neuschatz, & Goodwin, 1999, for related findings). Interestingly, however, there was no hypermnesia for the critical lures from the weakly associated lists in Experiment 1.

The item gain scores produced no significant effects for the list items, with all conditions producing relatively few gains. This result is consistent with the view that the list items received relatively little item-specific processing. For the critical lures, however, associative strength increased item gains, suggesting greater item-specific processing. The cumulative recall curves support this conclusion, with the critical lures from the strongly associated lists being the only items to produce relatively continuous recall throughout the cumulative recall period.

Unfortunately, the item loss scores suffered from a floor effect and produced no significant effects, limiting our

Figure 1. Mean cumulative recall percentages for both the list items and the critical lures as a function of associative strength in Experiment 1.
ability to make inferences about differences in relational processing. However, as evidenced by the substantially higher initial recall for the strongly associated items than for the weakly associated items, the cumulative recall curves suggest that the increased associative strength enhanced relational processing. Taken together, these results suggest that increasing associative strength not only resulted in an increase in relational processing of the list items but also produced an increase in item-specific processing of the critical lures. Apparently, the additional activation of the critical lures resulting from the more strongly associated list items is item specific in nature.

**EXPERIMENT 2**

In Experiment 2, in an effort to vary the type of processing performed on the list items, we manipulated the type of orienting task given to participants. In addition to the control group, which was given the standard memory instructions, a second group was required to form an item-specific image of each list item, and a third group was instructed to form relational images of list items interacting. Previous research has demonstrated that the item-specific imagery task promotes the encoding of item-specific information, whereas the relational imagery task leads to relational processing of the list items (see, e.g., Burns & Gold, 1999; Hodge & Otani, 1996).

It is well-known that orienting tasks that induce the same type of processing that would otherwise be performed have little effect on retention, whereas orienting tasks that require different types of processing generally improve performance (see, e.g., Einstein & Hunt, 1980; Hunt & Einstein, 1981). By requiring one group to perform relational processing and the other group to perform item-specific processing, we sought to determine the type of processing that the control group performed on the list items.

If Burns et al. (2006) are correct in their assertion that list items typically receive primarily relational processing, the relational imagery group should produce results similar to those of the control group, due to the repetitive processing. If, however, the list items normally receive item-specific processing, then we would expect the control and item-specific imagery groups to perform similarly due to processing redundancy.

Two previously published reports are relevant here. McCabe, Presmanes, Robertson, and Smith (2004) required participants to perform either a relational or an item-specific orienting task on DRM lists. They found that item-specific processing led to fewer critical lure false alarms than did relational processing. However, no control group was tested in their study, negating the opportunity to determine the type of processing typically given to list items in the DRM procedure. Smith and Hunt (1998, Experiment 3) compared a group given an item-specific orienting task (pleasantsness rating) with a group given standard instructions and found that the group that performed the item-specific task had fewer false memories. Smith and Hunt did not, however, include a group given a relational orienting task. Hence, it is possible that both groups given relational and item-specific orienting tasks will produce results different from those produced by the standard instructions group.

A second purpose of Experiment 2 was to further delineate the type of processing typically given to the critical lures. Burns et al. (2006) suggested that the critical lures receive more item-specific processing than do list items, and this item-specific advantage is responsible for the differences found in the item gain and cumulative recall scores. If this is the case, then requiring item-specific processing of list items should eliminate item-specific processing differences between the list items and the lures, thereby equating the item gain and cumulative recall scores. Hence, we predicted that whereas the item gain and cumulative recall differences between the list items and the critical lures would be obtained for the control and relational imagery groups, they would be eliminated for the item-specific imagery group.

**Method**

**Participants and Materials.** The 105 Union College students participating in the experiment were given either $4.00 or class credit in exchange for their participation. There were 36, 35, and 34 students tested in the control, relational imagery, and item-specific imagery groups, respectively. The same list of 120 words comprising 10 sublists was shown to all participants. The 10 sublists were taken from the lists used by Roediger and McDermott (1995). Each sublist consisted of the 12 strongest associates, with these associates blocked during presentation and presented in the order of strongest to weakest associative strength.

**Procedure.** The words were presented in sets of three on a computer monitor, one word to a row. Each set of words was presented for 12 sec. Participants in the control group were simply told that they would be shown the words and would later be asked to recall them. Participants in the item-specific imagery group were asked to form as vivid an image of each word as they could. To facilitate compliance with the instructions, these participants were asked to rate the vividness of each image and to write their ratings on a score sheet. Participants in the relational imagery group were asked to form as vivid an image as they could of the three words interacting with each other. They were also asked to rate the vividness of each image (of the three words interacting together) on a scale from 1 to 7 and to write their ratings on a score sheet.

Following list presentation, participants were given three successive recall tests, each lasting 4 min, with about a 1-min delay between tests. As in Experiment 1, they were required to draw a line under the last word recalled after each minute of recall.

**Results**

The mean recall percentages for each test for each of the three encoding conditions are shown in Table 2. A 3 (encoding group: control vs. item-specific imagery vs. relational imagery) × 3 (test number) × 2 (item type) ANOVA performed on the recall percentages produced a significant three-way interaction \( F(4,204) = 3.80, MS_e = 26.01 \). Separate 3 (encoding group) × 3 (test number) ANOVAs were performed on the recall percentages for each item type. For list items, both main effects and the interaction were highly significant (smallest \( F = 5.66 \) for the main effect of encoding group). Follow-up one-way ANOVAs comparing the three conditions on each recall test revealed that the conditions did not differ significantly on the first test \( F(2,102) = 1.84, MS_e = 55.77 \), but they did differ on the two latter tests \( F(2,102) = 6.55, MS_e = \)
64.99, and F(2,102) = 9.55, MS_e = 75.20, respectively]. Tukey tests revealed that on both the second and third tests, the item-specific imagery group produced higher list item recall than did the other two groups, which did not differ.

For the critical lures, the 3 \times 3 ANOVA revealed that recall differed significantly both across encoding groups [F(2,102) = 4.04, MS_e = 508.80] and across tests [F(2,102) = 35.95, MS_e = 54.55]. The interaction did not approach significance (F < 1). Collapsing across test number, a follow-up Tukey test showed that the item-specific imagery group recalled fewer critical lures than did the other two groups, which did not differ from each other.

The 3 (encoding group) \times 2 (test number) \times 2 (item type) ANOVA performed on the items-gained percentages showed that the only significant effects were the main effect of item type and the group \times item type interaction [F(1,102) = 9.10, MS_e = 30.36, and F(2,102) = 10.81, MS_e = 30.36, respectively]. Collapsing across test number, follow-up tests revealed that for both the control and the relational imagery groups, there were significantly more gains for the critical lures than for the list items [F(1,35) = 15.10, MS_e = 32.59, and F(1,34) = 11.30, MS_e = 21.88, respectively]. The opposite pattern occurred for the item-specific imagery group, although the items-gained advantage for the list items was not quite significant [F(1,33) = 3.13, MS_e = 36.74]. However, when the item gain percentages were calculated using the second formula (dividing the raw number of item gains by the total number of items presented minus the number recalled on the previous test), the item gains advantage for the list items was significant [F(1,33) = 6.77, MS_e = 56.45].

Additional Tukey tests revealed that the item-specific imagery group produced more list item gains than did the other two conditions (which did not differ), suggesting that the item-specific group encoded more item-specific information. For the critical lures, none of the groups differed significantly from each other. However, when we combined the control and relational imagery groups and compared them with the item-specific group, the difference approached significance [F(1,103) = 3.57, MS_e = 138.56, p = .06] and reached significance when we used the second formula for calculating item gain percentages [F(1,103) = 4.25, MS_e = 119.27].

The three-way ANOVA conducted on the item loss percentages failed to produce any significant effects (largest F = 2.81, MS_e = 10.85, for the group \times item type interaction), which was likely due to a floor effect since all conditions produced fewer than 4% losses.

Cumulative recall curves for all 12 min of recall (across the three tests) for each of the three encoding groups for both the list items and the critical lures are shown in Figure 2. For the list items, it is clear that the control and relational imagery groups produced similar curves, which approached asymptotic recall more rapidly than did the curve for the item-specific imagery group. As expected, these results suggest that relative to the control group, item-specific imagery enhanced item-specific processing, but relational imagery did not enhance relational processing. Second, for both the control and the relational imagery groups, cumulative recall was slower but steadier for the critical lures than for the list items. However, for the item-specific imagery group, the list items actually produced a steadier rate of recall than the critical lures did.

**Discussion**

We attempted to determine the type of processing typically given to the list items in the DRM procedure by using orienting tasks previously shown to induce either relational or item-specific processing. The relational and item-specific imagery tasks were apparently successful at inducing their respective types of processing: The relational imagery group produced list item cumulative recall curves indicative of primarily relational processing, whereas the item-specific imagery group produced list item cumulative recall curves suggestive of primarily item-specific processing. In addition, the item-specific group produced more list item gains than the other two groups did, indicating greater item-specific processing.

Comparing the two orienting-task groups with the control group revealed that for both the list items and the critical lures, the relational imagery and control groups

---

**Table 2**

Mean Performance Measures (in Percentages) and Standard Deviations for the List Items and Critical Lures for Each Encoding Group in Experiment 2

<table>
<thead>
<tr>
<th>Type of Item</th>
<th>Control Group</th>
<th>Relational Group</th>
<th>Item-Specific Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List</td>
<td>Critical</td>
<td>List</td>
</tr>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Test 2 recall</td>
<td>21.44</td>
<td>6.27</td>
<td>21.94</td>
</tr>
<tr>
<td>Test 3 recall</td>
<td>22.94</td>
<td>7.34</td>
<td>25.83</td>
</tr>
<tr>
<td>Test 2 gains</td>
<td>3.15</td>
<td>1.83</td>
<td>8.06</td>
</tr>
<tr>
<td>Test 3 gains</td>
<td>3.47</td>
<td>2.42</td>
<td>6.67</td>
</tr>
<tr>
<td>Test 2 losses</td>
<td>3.03</td>
<td>2.51</td>
<td>3.06</td>
</tr>
<tr>
<td>Test 3 losses</td>
<td>1.97</td>
<td>1.61</td>
<td>2.78</td>
</tr>
<tr>
<td>Approach to asymptote (λ)</td>
<td>0.55</td>
<td>0.19</td>
<td>0.55</td>
</tr>
<tr>
<td>Asymptote [M = 3]</td>
<td>25.15</td>
<td>33.04</td>
<td>28.32</td>
</tr>
</tbody>
</table>

Note—Test 3 recall does not always equal Test 2 recall plus Test 3 gains minus Test 3 losses, because items lost on the second test but gained on the third test were not included as item gains. **Approach to asymptote** refers to the estimated rate of approach to the asymptotic cumulative recall level using Equation 1 and is based on group means.
performed similarly. However, the item-specific imagery group produced results that differed considerably from those of the other two groups. These results support the contentions of Burns et al. (2006) that in the typical DRM procedure, list items receive mostly relational processing, whereas critical lures receive mostly item-specific processing. When the orienting task emphasized relational processing, there was little change in the recall of either the list items or the critical lures relative to the control group. However, when the orienting task encouraged item-specific processing, recall of the list items increased and recall of the critical lures decreased.

The item gain scores revealed that whereas critical lures produced significantly more item gains than did list items for the control and relational imagery groups, the item-specific imagery group produced a nonsignificant trend in the opposite direction. As predicted, requiring item-specific processing of the list items eliminated the item gain advantage for the critical lures, suggesting that both types of items received item-specific processing.

Although the effect produced was not large, the item-specific imagery group tended to produce fewer critical lure item gains than did the other two groups. Additionally, the critical lure cumulative recall curve for the item-specific imagery group suggested little use of item-specific processing relative to the control group's processing of the critical lures. These results support the view that the decrease in critical lure recall for the item-specific imagery group resulted from less item-specific processing of those lures.

It is possible that item-specific processing of the list items actually reduced the amount of item-specific information encoded with the critical lures. Alternatively, it may be that the critical lures received a substantial amount of item-specific processing, but the encoded item-specific information was not well utilized during retrieval. We favor the second explanation for two reasons. First, as described above, we contend that item-specific processing of the critical lures results from spreading activation via the list items.

There is little reason to believe that the amount of activation would decrease as a result of item-specific processing of the list items. Second, although critical lure recall was substantially lower following item-specific processing of the list items, the lures still produced a considerable number of item gains, suggesting that item-specific information was available (see, e.g., Burns, 1993; Klein et al., 1989). One explanation for why the encoded item-specific cues were not fully utilized by the item-specific group is that a considerable number of item-specific cues were also encoded with the list items, and these item-specific cues competed for access at time of retrieval. This idea is explored further in the General Discussion.

The recent proposal by Burns et al. (2006) that in the typical DRM procedure, more item-specific but less relational information is encoded with the critical lures than with the list items runs counter to extant theory. The prevailing view is that list items receive more item-specific processing than do critical lures. In fact, in most theoretical explications, it is this difference in item-specific processing that allows list items to be differentiated from the lures.

The results of the present experiments are consistent with Burns et al.'s (2006) contentions. Experiment 1 showed that increasing the associative strength between list items and lures increased the number of critical lures gained across successive tests, an index of item-specific processing. Moreover, the shapes of the cumulative recall curves for the critical lures indicated greater item-specific processing for the strongly associated lists. We conclude that increasing the associative strength between the list items and the critical lures, which presumably increases critical lure activation, enhances item-specific processing of the critical lures.

The relational imagery task used in Experiment 2, which presumably focused attention on relational information, had little effect on participants' recall of either the list items or the lures relative to the recall of participants in the control group. The item-specific imagery task, however, substantially increased recall of the list items and decreased recall of the critical lures. These results suggest that list items usually receive mostly relational processing, a conclusion bolstered by the cumulative recall analysis.

The typical item gain advantage for the critical lures relative to the list items was eliminated for the item-specific imagery group. In addition, the cumulative recall curves indicated that relative to the control group, the critical lures for the item-specific imagery group received little item-specific processing.

The findings of Burns et al. (2006) and of the present experiments suggest that critical lures in the DRM procedure typically receive substantial item-specific and little relational processing, a conclusion that is difficult to integrate with leading theories. Schacter and his colleagues (e.g., Dodson & Schacter, 2001; Schacter et al., 1999) suggested that reductions in false recollection following item-specific or distinctive processing of the list items.
resulted from participants’ demanding substantial item-specific detail before accepting an item as a previously experienced event. However, if the critical lures are rich in item-specific detail, as we contend, then item-specific information about the list items would not be diagnostic, and little reduction in false recollection would occur.

The activation/monitoring hypothesis (see, e.g., McDermott & Watson, 2001; Roediger et al., 2001) suggests that false memory reduction following item-specific processing of list items occurs during a monitoring process. The item-specific information is used to filter out the critical lures, which presumably are poor in item-specific information. Our results, however, suggest that the critical lures are rich in item-specific information.

The present results are also problematic for impoverished relational encoding theories (see Hege & Dodson, 2004; Schacter et al., 2001), which assume that increased item-specific processing of list items reduces false memories by diminishing relational processing of the critical lures. Our results provide little evidence that critical lures typically receive much relational processing or that increases in item-specific processing of the list items reduce relational processing of the lures.

One possibility is that critical lures receive primarily relational processing, but, since that relational information may not become part of the organized retrieval plan, it may behave like item-specific information at time of retrieval, improving recall later in the recall period. Burns et al. (2006) tested and rejected this hypothesis. They found that even when the critical items were physically presented in the DRM lists, they still produced more item gains, along with cumulative recall curves indicative of considerable item-specific processing. Hence, the item gain and cumulative recall results appear to reflect true differences in item-specific processing and are not simply an artifact of the critical lures’ not being physically present in the list.

Classifying Encoding Variables That Influence Illusory Recollection

As mentioned in the introduction, several encoding variables have been shown to influence false recollection. Theoretically, many of these variables can be classified into those likely to enhance item-specific processing of the list items and those that presumably enhance relational processing. Those variables likely to enhance item-specific processing include saying versus hearing the list items (Dodson & Schacter, 2001), presenting the items as pictures versus as words (see, e.g., Israel & Schacter, 1997), using instructions that require pleasantness ratings of the items versus standard instructions (Smith & Hunt, 1998), and presenting each item in a different type font versus in the same font (Arndt & Reder, 2003). For each of these variables, the manipulations presumed to increase item-specific processing not only had the expected effect of increasing veridical retention of the list items but also decreased false recollection of the critical lures.

Those variables that are likely to enhance relational processing of the list items include varying the presentation order of the associates (strongest to weakest or vice versa), the number of associates in the list (many or few), the strength of the associations between the list items and their respective critical lures (strong or weak), and the presentation order of the DRM lists (blocked by list or intermixed). These relational variables produce entirely different patterns of results. Specifically, conditions associated with more relational processing tend to produce higher levels of both veridical and false recollection (Gallo & Roediger, 2002; McDermott, 1996; McEvoy et al., 1999; Robinson & Roediger, 1997).

There are, of course, manipulations that cannot be easily categorized as enhancing either item-specific or relational processing. These variables include presentation rate, division of attention, number of repetitions of the list items, and number of study–test trials given. For example, dividing attention might be expected to decrease attention to item-specific details, or it might decrease the probability of forming associations between the list items, or both. These variables always affect veridical retention. Interestingly, however, results concerning recollection of the critical lures are far more equivocal (see, e.g., Benjamin, 2001; McDermott, 1995; McDermott & Watson, 2001; Payne, Lampinen, & Cordero, 1996; Pérez-Mata, Read, & Diges, 2002; Tussing & Greene, 1997, 1999).

Why does increasing item-specific processing of the list items decrease false recollection, whereas increasing relational processing increases false recollection? We suggest that increasing item-specific processing of the list items actually makes them more similar to the critical lures in terms of the amount of item-specific information processed. Thus, the critical lures no longer enjoy their item-specific (distinctive) advantage and, therefore, are less likely to be remembered. Increasing relational processing of the list items increases the amount of item-specific processing given to the critical lures, thereby increasing false recollection. This is precisely what we predicted in Experiment 1 when we increased the associative strength between list items and lures.

Perceptual Versus Conceptual Detail

Several studies have converged on the view that false memories for critical lures typically contain little structural or perceptual detail relative to list items (see, e.g., Arndt & Reder, 2003; Hicks & Starns, 2005). Since perceptual detail is a form of item-specific information, it is easy to assume that critical lures receive little item-specific processing. Perceptual detail, however, is only one type of item-specific information. A second type involves conceptual (i.e., meaningful or semantic) detail (see, e.g., McCabe et al., 2004; Rajaram, 1993). For example, encoding information about the pleasantness of an item and forming a mental image of a word’s referent represent two forms of conceptual detail that are likely to be item specific in nature. Conceptual detail would be item specific to the extent that the detail is not shared with the list items. Typically, conceptual detail facilitates retrieval far more than perceptual detail does, but perceptual detail does have some positive effect (see, e.g., Blaxton, 1989; Hunt & Mitchell, 1982; Hunt & Toth, 1990). It may be that the item-specific information encoded with the critical
lures is conceptual, whereas the item-specific information given to the list items is typically more perceptual. If this is the case, then it suggests that the effectiveness of reducing illusory recollection by requiring item-specific processing of the list items will depend on the nature of that item-specific information.

Although manipulations presumed to increase processing of both perceptual (see, e.g., Arndt & Reder, 2003) and conceptual (McCabe et al., 2004) detail of the list items have been shown to reduce false recollection, the reduction may have been achieved via separate mechanisms for the two types of detail. It is possible that perceptual detail primarily facilitates discrimination between the list items and the critical lures at time of retrieval (e.g., source monitoring), whereas conceptual detail may increase the accessibility of the list items. An increase in accessibility of the list items might decrease the accessibility of the critical lures (rather than decreasing their discriminability), if for no other reason than that the size of the functional search set would increase. The above speculation leads to the prediction that requiring both perceptual and conceptual item-specific processing of the list items would be particularly effective at reducing false memories, since both accessibility and discriminability are important components of accurate retrieval.

Some evidence for differential roles for perceptual and conceptual item-specific processing comes from false memory experiments using hybrid lists of semantic and phonological associates (see, e.g., Watson, Balota, & Roediger, 2003). In those experiments, DRM lists that included both words that were phonologically (perceptually) related and words that were semantically related to the critical lures produced markedly more false alarms than pure lists of only phonologically related or only semantically related words. One explanation for this overadditive effect is that phonological similarity decreased perceptual item-specific processing of the list items because several words, including the critical lure, shared the same perceptual features (see Watson et al. for a slightly different explanation). The lack of item-specific perceptual detail would hinder discrimination between the list items and the critical lures, resulting in an increase in false memories.

In recognition memory studies involving remember—know judgments, a typical finding is that falsely recognized critical lures often receive "remember" responses (see, e.g., Roediger & McDermott, 1995). We speculate that the remember judgments for critical lures may be guided by the relatively rich item-specific information processed with the lures.

Item-Based Versus Event-Based Distinctiveness

Recently, Hunt (2003) proposed the distinction between item-based and event-based distinctiveness. Event-based distinctive processing refers to the processing of differences between events, such as the encoding of the differences between two separate lists of words. Item-based distinctiveness refers to the processing of differences among items within an event, such as the differences among the words within a list. Hunt showed that manipulations designed to increase item-based distinctiveness (e.g., the pleasantness rating) improved both recall and number of hits in recognition. Manipulations intended to increase event-based distinctiveness (e.g., requiring that a different orienting task be performed for each list) had little effect on recall or hits but tended to decrease intrusions in free recall and false alarms in recognition.

To the extent that critical lures come to mind during study (see, e.g., Mather et al., 1997), performance in the DRM task may involve both event- and item-based distinctiveness. Event-based distinctiveness would be beneficial for discriminating between the two events of hearing list words and thinking about critical lures and would thus reduce false memories. Item-based distinctiveness would be necessary for discriminating among the presented list items, thereby improving list item recall. The results of Experiment 2 can be interpreted within this theoretical context. The processing performed by the relational imagery group was not item specific; thus, it would not lead to item-based distinctiveness. Moreover, this type of processing was not sufficiently different from the processing performed by the control group to improve event-based distinctiveness. Hence, neither recall of list items nor recall of critical lures was affected relative to the control group.

The processing performed by the item-specific imagery group was designed to increase item-based distinctiveness. Therefore, the observed increase in recall of the list items would be expected. The contribution of item-specific imagery to event-based distinctiveness is less clear. To the extent that the item-specific processing given to the list items was the same as that given to the critical lures (i.e., both were conceptual, item-specific processing), event-based distinctiveness would not have increased, resulting in no change in critical lure recall. However, if the item-specific processing given to the list items was different from that given to the critical lures (i.e., perceptual vs. conceptual), then event-based distinctiveness would have increased, thereby decreasing false memories. Our results are consistent with the latter possibility.

AUTHOR NOTE

This research was supported by a faculty research grant to D.B. from Union College. We thank Reed Hunt, Jason Watson, and an anonymous reviewer for their helpful comments on an earlier version of this manuscript. Correspondence concerning this article should be addressed to D. J. Burns, Department of Psychology, Union College, Schenectady, NY 12308 (e-mail: burnsD@union.edu).

REFERENCES


NOTES

1. The term *processing* is used here to mean altering a mental representation in such a way that components of that representation are altered or activated. We do not restrict our definition to processes that are necessarily deliberate or under conscious control. This definition allows us to discuss the processing of critical lures that were not physically presented and may not have come to mind but whose mental representations were altered.

2. We used group means, rather than individual scores, to estimate $\lambda$ and $n(=)$ because several participants recalled only 0, 1, or 2 critical lures. With so few items recalled, either the estimates were not calculable or the data were fit poorly by Equation 1.

(Manuscript received July 10, 2006; revision accepted for publication January 5, 2007.)