Original Article

QJEP

Quarterly Journal of Experimental Psychology
2019, Vol. 72(6) 1493–1506

DOI: 10.1177/1747021818801427
qjep.sagepub.com
SAGE

Item-specific and relational processing both improve recall accuracy in the DRM paradigm

Mark J Huff[1] and Glen E Bodner[2]

Abstract

Using the Deese–Roediger–McDermott (DRM) paradigm, Huff and Bodner found that both item-specific and relational variants of a task improved correct recognition, but only the item-specific variants reduced false recognition, relative to a read-control condition. Here, we examined the outcome pattern when memory was tested using free recall, using the same item-specific versus relational task variants across three experiments as our previous study (processing instructions, pleasantness ratings, anagram generation). The outcome pattern in recall was similar to recognition, except relational processing at study actually reduced the DRM illusion, though not as much as item-specific processing. To reconcile this task difference, we suggest that the memory information laid down during relational encoding enhances the familiarity of the critical items at test. To the extent that familiarity is used less as a basis for responding in free recall than in recognition, relational processing ironically reduces rather than increases the DRM illusion in recall.

Keywords

Item-specific processing; relational processing; distinctiveness; free recall

Received: 10 January 2018; revised: 31 March 2018; accepted: 2 August 2018

Distinctive processing refers to the processing of "difference within a context of similarity" (Hunt, 2006, p. 12; see also Schmidt, 1991). In memory, the processing of distinctive features often yields a memorial advantage. At the perceptual level, distinctive processing can improve memory accuracy through such manipulations as studying pictures (vs. words) of list items (Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999), using orthographically bizarre versus typical words (Hunt & Elliot, 1980; McDaniel, Cahill, & Bugg, 2016), and studying items in isolation (Kelley & Nairne, 2001; von Restorff, 1933). Relatedly, distinctive processing can be induced through encoding tasks such as those that encourage item-specific processing, such as pleasantness ratings or generation (Huff & Bodner, 2013; Hunt, Smith, & Dunlap, 2011). In the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), which is our focus, item-specific processing has been found to enhance overall memory accuracy by simultaneously increasing correct memory and reducing false memory (e.g., Huff & Bodner, 2013; McCabe, Presmanes, Robertson, & Smith, 2004). The present study examines whether the pattern of influence of a given type of processing on memory accuracy in the DRM paradigm differs when memory is tested via free recall.

The standard DRM paradigm presents participants with lists of associates (e.g., *bed, tired, rest*) that converge upon a single non-studied critical item (CI; for example, *sleep*). When tested, false memory for CIs is high: False recall rates approach 50% (Roediger & McDermott, 1995) and false recognition of CIs can meet or exceed hit rates (see Gallo, 2006 for review; Lampinen, Neuschatz, & Payne, 1999). Furthermore, participants report conscious recollection of CIs as occurring on a studied list as evidenced by *remember* responses at test, even though CIs were internally generated (Payne, Elie, Blackwell, & Neuschatz, 1996).

[1]The University of Southern Mississippi, Hattiesburg, MS, USA
[2]College of Education, Psychology and Social Work, Flinders University, Adelaide SA, Australia

Corresponding author:
Mark J Huff, School of Psychology, The University of Southern Mississippi, Hattiesburg, MS 39406, USA.
Email: mark.huff@usm.edu

The DRM illusion is typically accounted for through either a gist-extraction or an activation-monitoring process. According to fuzzy-trace theory (Brainerd & Reyna, 1990, 1998), studying a list of words encodes both a verbatim and gist memory representation. The verbatim representation contains the specific list item or the contextual details of the item, and the gist representation contains the overall thematic meaning of the list. As the CI does not contain a verbatim trace (because it was not studied), the DRM illusion is suggested to occur due to a stored gist representation, which is strengthened through the thematic consistency of the list items (Brainerd & Reyna, 2002). Separately, activation-monitoring theory (Roediger, Balota, & Watson, 2001) posits that the DRM illusion reflects a two-stage process in which the CI is activated implicitly through a spreading-activation process at encoding (Collins & Loftus, 1975), followed by a monitoring failure to reject this activation at test. In the standard DRM paradigm, both accounts are confounded as list items are thematically and associatively related. Recently, however, paradigms examining various levels of thematic and associative relations have supported both accounts (Huff, Coane, Hutchison, Grasser, & Blais, 2012; Huff, McNabb, & Hutchison, 2015; Hutchison & Balota, 2005), suggesting that both gist extraction and associative activation contribute to the DRM illusion.

Given the robustness of the DRM illusion, researchers have often sought methods with which to reduce it. Successful methods include repeated study presentations (Benjamin, 2001; McDermott, 1996), instructional warnings presented before study and/or test (Gallo, Roediger, & McDermott, 2001; McCabe & Smith, 2002; Neuschatz, Payne, Lampinen, & Toglia, 2001), and requiring source monitoring of the origins of one's memories at test (Johnson, Hashtroudi, & Lindsay, 1993; Multhaup & Conner, 2002; although see Hicks & Marsh, 2001, for exception). Manipulations that induce distinctive encoding of the list items have also been successful at reducing the illusion, including picture presentations (Israel & Schacter, 1997), unique fonts (Arndt & Reder, 2003), mental imagery (Bodner, Huff, Lamontagne, & Azad, 2017; Foley, Wozniak, & Gillum, 2006; Gunter, Bodner, & Azad, 2007; Oliver, Bays, & Zabrucky, 2016; Robin, 2010), pleasantness ratings (Gunter et al., 2007; Huff & Bodner, 2013), and anagram generation (Bodner et al., 2017; Gunter et al., 2007; Huff & Bodner, 2013; McCabe & Smith, 2006). Some of these distinctive manipulations, such as generation, can induce a *mirror effect* (Glanzer & Adams, 1990)—an increase in correct memory coupled with a decrease in false memory—relative to a read-only control task.

Two loci for the benefits of distinctive processing have been posited: enhanced encoding and enhanced memory monitoring at test (see Huff, Bodner, & Fawcett, 2015, for a review). According to the *impoverished relational encoding* account (Hege & Dodson, 2004; Hockley & Cristi, 1996), distinctive processing operates at study by disrupting the thematic consistency of the list (Brainerd & Reyna, 2002) or by reducing the implicit associations shared between list items and the CI (Roediger et al., 2001). However, distinctive processing can also lead participants to adopt a global-decision strategy at test termed a *distinctiveness heuristic*. This heuristic refers to a test-based monitoring process wherein memory items are reported only when accompanied by the recollection of distinctive details from the encoding task. Recollection of these details provides diagnostic evidence that the item was studied, whereas their absence can be used to reject items through a recall-to-reject disqualifying monitoring process (Gallo, 2004).

In a recent series of studies, we have used signal-detection indices to demonstrate that both mechanisms can contribute to the memory improvements gained via distinctive processing, at least when memory is tested via recognition (Bodner et al., 2017; Gunter et al., 2007; Huff & Bodner, 2013; Huff, Bodner et al., 2015; for a similar account, see Hunt & Smith, 2014). Here, we illustrate this pattern using the Huff and Bodner study, which served as the template for our current study. In each of the three experiments, we compared the effects of item-specific versus relational encoding on correct and false recognition relative to a read-only control group. We then applied our signal-detection approach to measure the effects of each type of processing on encoding (memory information, as indexed by d') and retrieval (memory monitoring via a distinctiveness heuristic, as indexed by lambda). At study, the item-specific groups focused on the unique features of each of the DRM list items, whereas the relational groups focused on the shared characteristics of the list items (Hunt & Einstein, 1981). Across experiments, item-specific and relational variants were created via different processing instructions, or were embedded in either a pleasantness rating task or an anagram generation task.

As highlighted in Figure 1 (top panel), across experiments, Huff and Bodner (2013) found that both item-specific and relational variants improved correct recognition relative to the read-control groups. This improvement was due to a combination of more information having been encoded about the list items (as shown by a higher memory information index) and an increase in memory monitoring at test (as shown by a higher memory monitoring index). As Figure 1 also illustrates, item-specific processing also reduced false recognition, by both reducing the amount of encoded memory information for CIs and increasing monitoring at test, relative to the read groups. Relational processing, in contrast, did not reduce the amount of information encoded about the CIs, but it did increase monitoring at test (as much as for item-specific processing, as shown in the right panel of Figure 1). Critically, the increase in monitoring following relational processing was ineffective at

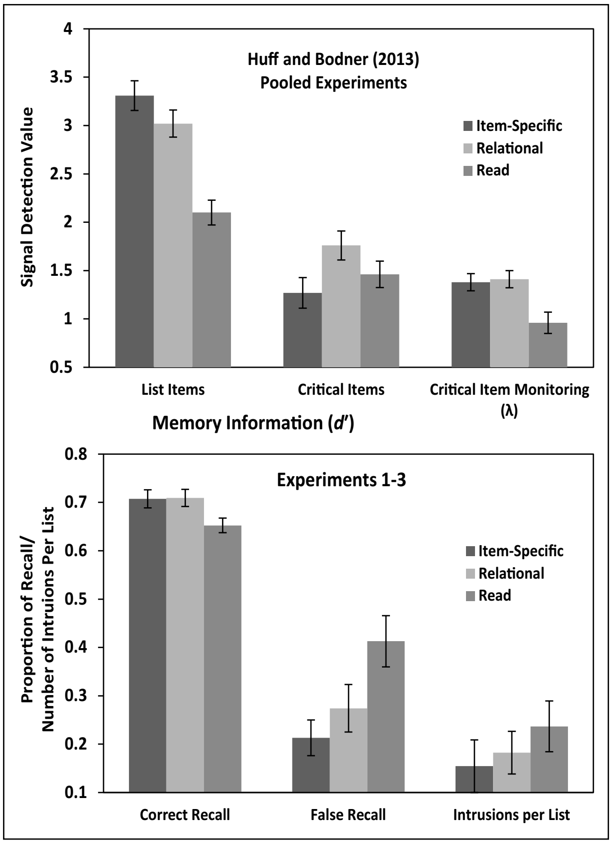

Figure 1. Pooled recognition data from Experiments 1 to 3 from Huff and Bodner's (2013; top) study, and pooled recall data from Experiments 1 through 3 of the current study (bottom).

reducing false recognition because information about the CIs had already been encoded. Thus, there was evidence for elevated monitoring akin to a distinctiveness heuristic in both the item-specific and relational groups, but it was not effective for the relational groups. In general, whether a given encoding task improved memory accuracy depended less on the type of task and more on whether it was set up to primarily induce item-specific or relational processing.

The signal-detection approach thus provides valuable insights into the effects of different encoding manipulations and strategies on recognition accuracy. However, Huff and Bodner (2013) noted two limitations of their signal-detection approach. First, the signal-detection indices are strictly quantitative—they capture the *amount* of memory information and the *amount* of monitoring—and cannot identify qualitative differences in specific types of information that participants encode or monitor for at test. Second, the approach cannot be applied in tasks other than recognition, such as free recall.

The primary goal of our study was to investigate whether item-specific and relational processing in the DRM paradigm yield similar or different outcomes when memory is tested via free recall rather than recognition. There have been several demonstrations that distinctive encoding can reduce the DRM illusion in recall, including picture presentation (Oliver et al., 2016; Smith, Hunt, & Dunlap, 2015), pleasantness ratings (Hunt et al., 2011), and anagram generation (McCabe & Smith, 2006), though few studies have compared distinctive encoding relative to both relational encoding and a read control. In a notable exception, Thomas and Sommers (2005; Experiment 2) embedded DRM list words in sentences that either converged or diverged with the list theme which were then compared with a read task in younger and older adults. Consistent with item-specific encoding effects, the DRM illusion was reduced for divergent sentences relative to convergent sentences and reading, which, in turn, were equivalent in both age groups. However, correct recall was lower for both sentence conditions relative to read, indicating that divergent sentences did not improve correct memory, contrary to what is usually found with item-specific tasks in recognition. Separately, Burns, Jenkins, and Dean (2007) found that item-specific instructions produced a mirror effect in recall relative to relational and read conditions. However, Butler, McDaniel, McCabe, and Dornburg (2010) found both a reduction in correct recall and an increase in DRM false recall compared with reading using a similar set of item-specific instructions in older adults. Therefore, although item-specific and relational processing effects are relatively consistent in recognition, the pattern in recall has been less consistent. Our experiments therefore aimed to provide additional tests of the reliability of item-specific and relational processing effects on correct and false recall.

There are reasons to expect that the effects of item-specific versus relational processing may show a different outcome pattern in recall than they did in recognition. Typical implementations of distinctive encoding tasks emphasise item-specific processing and primarily enhance recollection (e.g., Yonelinas, 2002). In contrast, relational versions of those same tasks, which focus on thematic or associative processing of the list items (e.g., Brainerd & Reyna, 2002; Roediger et al., 2001), primarily enhance CI familiarity. Furthermore, it is generally assumed that recall is more monitoring-heavy and recollection-based and, in turn, is less automatic and familiarity-based than recognition (Koriat & Goldsmith, 1996; Mandler, 1980; Roediger & Guynn, 1996; Yonelinas, 2002), given that participants must self-generate memory candidate items in recall. A counterintuitive prediction follows, namely, that relational processing of DRM lists at study might improve free-recall accuracy relative to a standard read-control task. Specifically, relational processing should lead the CIs to feel familiar when they come to mind at test, but if familiarity is not the primary basis for outputting items in recall, then participants might not output them. If so, relational processing might not increase false memory relative to item-specific processing in recall, unlike the case when recognition memory is tested.

To accomplish our primary goal, we report three free-recall experiments that compared the effects of the same item-specific and relational processing task variants (relative to a read-control group) used by Huff and Bodner (2013) in recognition: processing instructions (Experiment 1), pleasantness ratings (Experiment 2), and anagram generation (Experiment 3). In each of the experiments, the item-specific and relational task variants were designed to encourage the processing of unique or related features, respectively. Although pleasantness ratings and generation tasks are traditionally used as item-specific tasks (see Gunter et al., 2007; Hunt & Einstein, 1981), Huff and Bodner demonstrated that the type of processing recruited is contingent upon structure of these tasks and not the task itself. We expected item-specific variants to elicit a mirror effect pattern relative to the read group. The relational variants were similarly expected to increase correct recall due to the elaborative processing involved, but unlike in recognition, we expected that relational processing might also work to reduce false recall, based on the logic laid out above.

In the absence of our signal-detection approach, a secondary goal of our study was to examine whether measures of recall dynamics might be useful for identifying the encoding and monitoring processes used in free recall. As an example, Burns et al. (2007) found differences in the recall rates between item-specific, relational, and read instructions using cumulative-recall curves plotted over three consecutive recall tests. These differences demonstrate that aspects of the recall process beyond mean recall rates can be informative regarding how processing instructions at encoding affect item output during recall.

In our current experiments, we report three metrics (for others, see the Online Supplementary Material). First, we measured whether item-specific and relational processing strengthen memory for list items from similar serial positions relative to reading. Item-specific processing has been shown to diminish primacy effects but enhance recall from the middle positions and the recency portion of a list relative to reading (Seiler & Engelkamp, 2003). We expected that item-specific encoding would produce a similar pattern, but we did not know whether the same would be true of relational processing. However, given that both types of processing can be considered "deep" in terms of level of processing, their serial position curves might well be similar. Second, we measured whether the test output position of falsely recalled CIs differed across groups. We note that this analysis has previously been reported by Roediger and McDermott (1995) but, to our knowledge, has not been used to examine processing effects on false recall. If the three groups monitor their memory in qualitatively different ways during recall, then the output position for CIs could differ. Of course, differences in these recall dynamics could reflect contributions of either encoding or retrieval process, but nevertheless they would provide useful information about how processing types influence correct and false recall. Finally, we also compared the rate of extra-list intrusions for the two task variants relative to reading. If item-specific and relational task variants lead to more stringent memory monitoring at test, we might observe a reduction in extra-list intrusions relative to the read group.

Experiment 1: processing instructions

In Experiment 1, groups receiving either item-specific or relational processing instructions for a set of DRM lists were compared with a read-only control group. Both types of instruction were expected to improve correct recall relative to reading. Item-specific instructions were further expected to reduce false recall relative to reading. Unlike in recognition, we expected that relational instructions might reduce false recall relative to reading. This expectation follows because relational processing should work to increase the familiarity of the CIs, and recall is less reliant on familiarity. Extra-list intrusions and recall dynamics were also analysed to provide novel information regarding the locus of the effects of these two processing variants on recall accuracy. We also explored whether extra-list intrusions provide a potentially sensitive indirect method of gauging monitoring processes at test, which, based on Huff and Bodner (2013), are expected to be used more by the item-specific and relational groups relative to the read-control group. We expected that extra-list intrusion rates would be low when monitoring processes were elevated.

Method

Participants. University of Calgary undergraduates participated for psychology course credit and were randomly assigned to either the item-specific, relational, or read instruction group (36 per group).

Materials. The 12 DRM lists yielding the highest rates of CI recognition in Huff and Bodner (2013) were divided into two counterbalanced sets of six lists. List set was assigned to participants at random. The six lists were then randomly ordered for each participant. Each list contained 12 items at least four letters in length, presented in descending order of backward-associative strength.

Procedure. Participants were tested individually via computer. They were informed that they would see a series of lists consisting of words displayed individually and would need to be read the items on the lists aloud. The item-specific group was further instructed that for each word, they should "think of a unique characteristic of each word that differentiates it from the other words on the list." The relational group was further instructed to "concentrate on what the words in that list have in common and associate them

Table 1. Mean (*SD*) proportion of recall of correctly studied list items, critical items, and number of intrusions recalled per list as a function of read, item-specific, and relational groups in Experiments 1 through 3.

	Processing groups		
	Item specific	Relational	Read
Experiment 1: Processing instructions			
List items	0.71 (0.09)	0.73 (0.10)	0.64 (0.08)
Critical items	0.26 (0.22)	0.37 (0.27)	0.44 (0.28)
Intrusions per list	0.21 (0.30)	0.26 (0.27)	0.29 (0.31)
Experiment 2: Pleasantness ratings			
List items	0.75 (0.08)	0.72 (0.09)	0.67 (0.07)
Critical items	0.21 (0.17)	0.22 (0.24)	0.40 (0.28)
Intrusions per list	0.06 (0.13)	0.12 (0.19)	0.16 (0.19)
Experiment 3: Anagram generation			
List items	0.66 (0.11)	0.68 (0.09)	0.64 (0.08)
Critical items	0.16 (0.19)	0.23 (0.25)	0.40 (0.28)
Intrusions per list	0.19 (0.37)	0.16 (0.21)	0.26 (0.30)

SD: standard deviation.

together." These encoding instructions were taken verbatim from Huff and Bodner (2013). Participants then received a practice list (the "cold" DRM list). The read group simply read the list aloud. The item-specific and relational groups verbalised their encoding with feedback, and further instruction was provided by the experimenter, when necessary. These groups were then informed that they could perform their encoding task silently for the remaining lists. Participants were instructed that for each word, they were to perform their encoding task silently and then read the word aloud. The experimenter advanced to the next word after each word was read aloud.

Participants next studied the first critical list, which was self-paced. Immediately following the study of each list, participants were prompted to complete a free-recall task where they were given a numbered sheet of paper and asked to write down as many words as they could remember from that list without cost for misspellings, for 2 min, on a sheet of paper corresponding to that list. At the end of 2 min, the participant turned the sheet over to the next recall sheet, and then the experimenter pressed a key to present the next study list. This sequence continued for six study/free-recall cycles. The experiment took about 30 min.

Results

We set a $p < .05$ significance level for analyses and provide effect sizes for reliable and marginal analyses of variance (ANOVAs) effects (η_p^2) and *t* tests (Cohen's *d*).

Recall. Table 1 presents the mean proportion of recalled list items (correct recall), CIs (false recall), and the average number of recalled extra-list intrusions per list (not including the CI), for each group. Correct recall of list items differed across the three groups, $F(2, 105) = 9.12$, mean squared error (MSE) = .01, $\eta_p^2 = .15$. As expected, relative to the read group, correct recall was greater following item-specific encoding (.64 vs. .71), $t(70) = 3.39$, standard error of the mean (SEM) = .01, $d = .81$, and relational encoding (.64 vs. .73), $t(70) = 4.06$, SEM = .01, $d = .97$. Item-specific and relational instructions led to similar levels of correct recall, $t < 1$.

False recall of CIs also differed across groups, $F(2, 105) = 4.43$, MSE = .07, $\eta_p^2 = .08$. Relative to the read group, false recall was reduced by item-specific encoding (.44 vs. .26), $t(70) = 3.03$, SEM = .04, $d = .72$, but the reduction found following relational encoding was not significant (.44 vs. .37), $t(70) = 1.13$, SEM = .07, $p = .26$. Moreover, false recall was only marginally lower following item-specific encoding than following relational encoding (.26 vs. .37), $t(70) = 1.85$, SEM = .04, $p = .07$, $d = .44$.

Although we had expected to find lower rates of extra-list intrusions following item-specific and relational processing relative to the read group, indicative of greater memory monitoring, such intrusions were rare and their rate did not differ across the groups, $F < 1$.

Recall dynamics. Correct recall was further analysed by comparing the effects of item-specific and relational encoding on the shape of serial position curves (see Crowder, 1976, for review). The serial recall curves shown in Figure 2 (top panel for Experiment 1) were computed by taking the average number of items recalled as a function of their serial positioned order in the studied list, smoothed over three adjacent serial positions. All groups showed large primacy and recency effects typical of immediate recall (see Lewandowsky & Murdock, 1989, for review). Importantly, encoding task differences emerged in the middle serial positions. Correct recall in Serial Positions 4 to 9 was higher (and equivalent) for the item-specific and relational groups relative to the read group. These patterns were confirmed by effects of group, $F(2, 105) = 9.15$, MSE = .11, $\eta_p^2 = .15$; serial position, $F(11, 1155) = 142.97$, MSE = .01, $\eta_p^2 = .58$; and an interaction, $F(22, 1155) = 3.46$, MSE = .01, $\eta_p^2 = .06$. Item-specific and relational instructions thus enhanced recall for items from the middle of the DRM lists, rather than increasing primacy and recency effects.

The CIs were generally recalled quite late during recall, as initially reported by Roediger and McDermott (1995). Their average output position was 6.91 out of an average of 9.06 total items (including intrusions). However, the mean output position of CIs was similar across the item-specific ($M = 6.46$, standard deviation [SD] = 2.36), relational ($M = 7.22$, $SD = 2.26$), and read ($M = 7.04$, $SD = 1.76$) groups, $F = 1.00$. Thus, item-specific and relational processing did not affect the output position of CIs during recall.

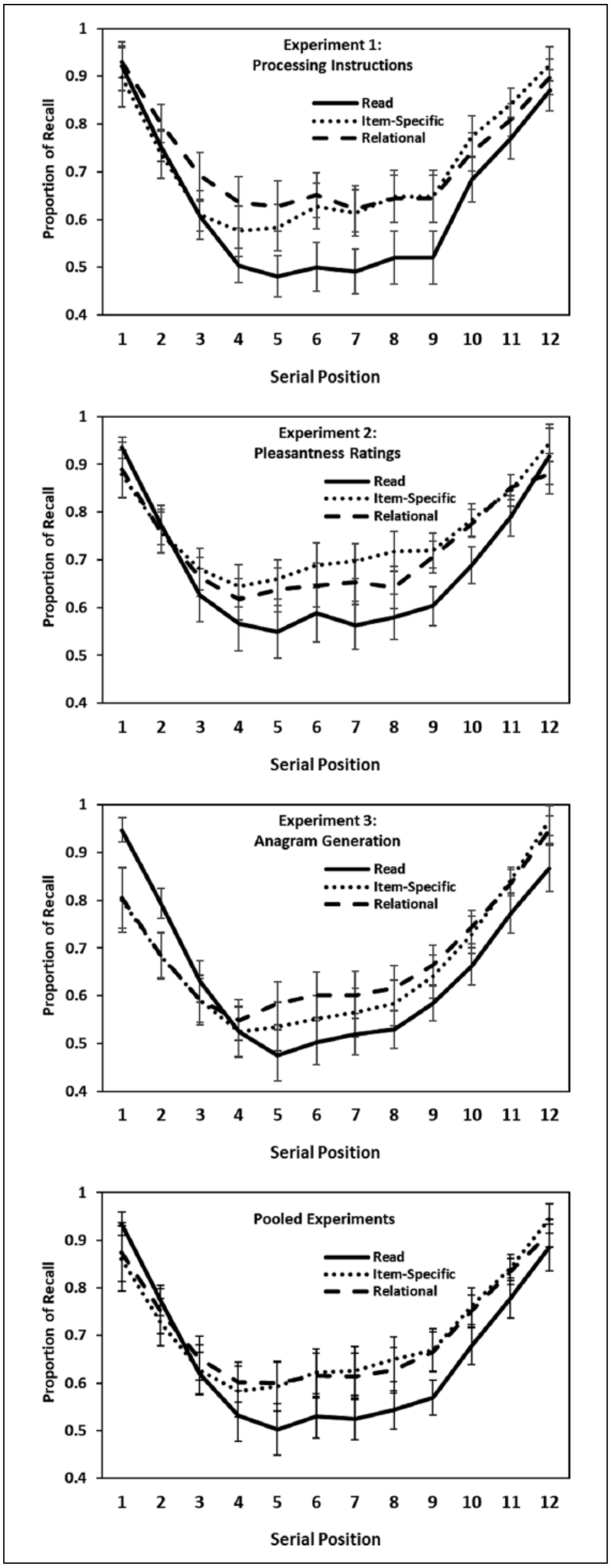

Figure 2. Serial position curves as a function of group in Experiments 1 through 3 (top 3 panels) and pooled data from Experiments 1 through 3 (bottom panel). Data were smoothed over three adjacent serial positions. Error bars are the 95% confidence intervals for the means.

Discussion

Experiment 1 investigated the influence of item-specific and relational processing instructions on recall accuracy in the DRM paradigm. Both processing types increased correct recall relative to reading, but only item-specific processing reliably reduced the DRM illusion. This pattern in recall replicates the pattern in recognition reported by Huff and Bodner (2013). We had reasoned that relational processing might also have successfully reduced the DRM illusion, if recall relies less on familiarity than recognition and if relational processing enhances familiarity of the CIs. That pattern was not in evidence. Moreover, we did not find evidence that monitoring was more stringent in the instruction groups: Rates of extra-list intrusions were similar across groups, though intrusions were rare across groups. Finally, our analyses of recall dynamics revealed that both types of instruction enhanced correct recall primarily of items from the middle positions. They also revealed that the CIs were typically output late in recall and this pattern did not differ across groups. Thus, there were no task dissociations in the effects of item-specific and relational processing, apart from false recall of CIs being marginally different. We hold off on further interpretation of the Experiment 1 findings until after we report Experiments 2 and 3 as well as a pooled analysis.

Experiment 2: pleasantness ratings

Experiment 2 compared the influences of item-specific and relational variants of a pleasantness task on recall with a read-control group. Pleasantness ratings have traditionally been used as an item-specific processing task (Burns & Schoff, 1998; Huff & Bodner, 2014; Hunt & Einstein, 1981; Toglia, Neuschatz, & Goodwin, 1999). In the DRM paradigm, this variant increases correct memory and reduces false memory relative to a read-control task (Huff & Bodner, 2013; Hunt et al., 2011). However, Huff and Bodner reported that both item-specific and relational variants of this task increased correct recognition, but only the item-specific variant reduced false recognition. In Experiment 2, both variants of the pleasantness task were expected to increase correct recall relative to reading, and the item-specific variant was also expected to reduce false recall of the CIs. Of major interest was whether the relational variant also reduces the DRM illusion, despite that outcome not having emerged statistically in Experiment 1.

Method

Participants. Additional students from the Experiment 1 pool were randomly assigned to the read, item-specific pleasantness, or relational pleasantness groups (36 per group).

Materials and procedure. The materials and procedure were identical to Experiment 1 with the following modifications. The pleasantness instructions were taken verbatim from Huff and Bodner (2013). The item-specific pleasantness group was instructed to "think of a unique reason why each list word is pleasant or not." Item-specific participants then judged each word as pleasant or not by pressing a labelled "yes" or "no" key on a response box. The relational pleasantness group was instructed to "rate the pleasantness of each word relative to the other words on the list using a 1-7 scale, with 7 representing most pleasant." They were further instructed to "rate the pleasantness of the first word individually, but beginning with the second word, rate its pleasantness relative to the first word, the third word relative to the first word, and so on." Relational participants made their ratings using keys on a response box labelled 1 to 7.

Results

Recall. Table 1 reports the Experiment 2 means. Correct recall differed across groups, $F(2, 105)=7.68$, MSE = .01, $\eta_p^2=.13$. Relative to the read group, correct recall was greater following item-specific pleasantness (.67 vs. .75), $t(70)=4.06$, SEM = .01, $d=.98$, and relational pleasantness (.67 vs. .72), $t(70)=2.53$, SEM = .01, $d=.60$. Correct recall was similar for the item-specific and relational variants, $t(70)=1.23$, SEM = .01, $p=.23$. Thus, as in Experiment 1, both variants improved correct recall relative to the read group.

False recall of the CIs also differed across encoding groups, $F(2, 105)=7.20$, MSE = .05, $\eta_p^2=.12$. False recall was lower following item-specific processing than after reading (.21 vs. .40), $t(70)=3.41$, SEM = .04, $d=.81$. Unlike in Experiment 1, false recall was also lower after relational processing than after reading (.22 vs. .40), $t(70)=2.89$, SEM = .04. False recall was equivalent between the two pleasantness variants, $t<1$.

The mean number of extra-list intrusions reported per list differed marginally across groups, $F(2, 105)=2.84$, MSE = .03, $p=.06$, $\eta_p^2=.05$. Fewer intrusions were reported in the item-specific pleasantness group than in the read group (.15 vs. .06), $t(70)=2.51$, SEM = .03, $d=.60$, whereas the intrusion rate was similar in the read and relational groups (.12 vs. .16), $t<1$. Intrusion rates did not differ reliably for the item-specific and relational groups (.06 vs. .12), $t(70)=1.56$, SEM = .03, $p=.12$. Again, intrusion rates were low across groups.

Recall dynamics. As shown in Figure 1 (middle panel), strong primacy and recency effects were again found and group differences emerged only in the middle serial positions. Specifically, correct recall was greater for the item-specific and relational groups than for the read group in Serial Positions 4 to 10 and was otherwise equivalent. Moreover, the curves for the item-specific and relational pleasantness groups were similar across serial positions. These effects were confirmed by main effects of encoding group, $F(2, 105)=7.10$, MSE = .01, $\eta_p^2=.12$; serial position, $F(11, 1155)=121.89$, MSE = .01, $\eta_p^2=.54$; and a significant interaction, $F(22, 1155)=3.85$, MSE = .01, $\eta_p^2=.07$. Thus, as in Experiment 1, item-specific and relational pleasantness tasks enhanced recall of items from the middle serial positions of the lists.

The mean output position for CIs was 7.38 out of an average of 9.10 items and did not differ significantly across the item-specific ($M=7.45$, $SD=2.07$), relational ($M=7.79$, $SD=2.07$), and read groups ($M=7.02$, $SD=2.10$), $F<1$.

Discussion

The results of Experiment 2 largely replicated Experiment 1, with one key exception: Here, relational processing reliably reduced false recall of CIs relative to reading. Indeed, the reduction was as large after relational processing as it was after item-specific processing. Thus, relational processing reduced the DRM illusion in free recall, whereas it did not do so in recognition in Huff and Bodner (2013). We ascribe this outcome to relational processing primarily enhancing the familiarity of the CIs, and the recall task being less sensitive to familiarity as a basis for responding. Experiment 3 sought additional evidence for this pattern.

Experiment 3: anagram generation

Experiment 3 compared the effects of item-specific and relational variants of an anagram generation task. Item-specific versions of the generation task have been shown to reduce the DRM illusion through a reduction in encoded memory information for the CI and/or an increase in memory monitoring at test (i.e., a test-based distinctiveness heuristic; Bodner et al., 2017; Gunter et al., 2007; Huff & Bodner, 2013; McCabe & Smith, 2006). In contrast, Huff and Bodner found that a relational version of the generation task failed to reduce the DRM illusion in recognition. We manipulated whether the generation task emphasised item-specific or relational processing, and then measured the effects on correct and false recall relative to a read group. We were particularly interested in whether relational processing would reduce the DRM illusion in recall, as was the case in Experiment 2 but not reliably so in Experiment 1.

Participants

Additional students from the Experiment 1 pool were randomly assigned to the item-specific generation ($N=36$), relational generation ($N=35$), and read ($N=35$) groups.

Materials and procedure

Following Huff and Bodner (2013), anagrams of the Experiment 1 lists were constructed by exchanging the first and third letters (e.g., *rahd* for *hard*) or second and fourth letters (e.g., *plliow* for *pillow*). Each list contained six anagrams of each type, assigned to list words at random once, and then fixed for all participants. For the relational generation group, each anagram was presented alongside an associatively related cued word (e.g., FEATHER—*rahd*, FEATHER—*liplow*). The same cue word was presented with all 12 words on a given list. The cue word was the word with the highest backward-associative strength from the list word to the CI that was not already on the list based on the Nelson, McEvoy, and Schreiber (1999) word-association norms.

The procedure was identical to Experiment 1 with the following modifications. The item-specific generation group was instructed to solve each anagram and to say the solution aloud. The relational generation group was similarly instructed, and they were further instructed to use the related cue word to help them find the solution. If an anagram could not be solved within a few seconds, the experimenter verbally provided participants with the first letter of the solution. Participants who still could not solve the anagram were told to say "pass." Accordingly, the experimenter coded the participant's response as "correct," "hint," or "pass." On "pass" trials, participants were told the correct solution.

Results

The mean anagram completion rates were 95.9% and 97.2% in the item-specific and relational groups, respectively, so our analyses were not conditionalised on correct generation.

Recall. Table 1 reports the means for Experiment 3. Unlike in Experiments 1 and 2, correct recall did not differ across the item-specific generation (.66), relational generation (.68), and read (.64) groups, $F < 1$. The absence of a generation effect was unexpected given its occurrence in recognition in Huff and Bodner (2013; see also Bodner et al., 2017; Gunter et al., 2007). However, the recall dynamics reported below provide evidence of an effect of generation.

False recall of CIs differed across groups, $F(2, 103) = 8.70$, MSE = .06, $\eta_p^2 = .15$. Replicating Experiment 2, relative to the read group, false recall was much lower in both the item-specific generation group (.40 vs. .16), $t(69) = 4.12$, SEM = .04, $d = .99$, and the relational generation group (.40 vs. .23), $t(68) = 2.63$, SEM = .04, $d = .64$. False recall was similar for the item-specific and relational variants (.16 vs. .23), $t(69) = 1.27$, SEM = .04, $p = .21$.

The rate of extra-list intrusions did not differ across the item-specific (.19), relational (.16), and read (.26) groups, $F(2, 103) = 1.03$, MSE = .09, $p = .23$.

Recall dynamics. As shown in Figure 1 (bottom panel), primacy and recency effects were again found as noted by an effect of serial position, $F(11, 1144) = 185.90$, MSE = .01, $\eta_p^2 = .64$. The main effect of encoding group was not significant, $F < 1$, but group and serial position interacted, $F(22, 1144) = 6.00$, MSE = .01, $\eta_p^2 = .10$. As shown in Figure 1, the read group showed a more robust primacy effect than the two generation groups (for a similar result, see Seiler & Engelkamp, 2003). In contrast, the generation groups showed greater recall than the read group from Position 5 onward, including for the recency portion of the curve. Thus, generation did enhance correct recall, but not from the primacy portion of the list.

The mean output position of CIs was 6.64 out of an average of 8.40 items and was again equivalent across item-specific ($M = 6.42$, $SD = 2.03$), relational ($M = 7.25$, $SD = 1.45$), and read ($M = 6.34$, $SD = 1.77$) groups, $F(2, 65) = 1.84$, MSE = 3.08, $p = .17$. Item-specific and relational processing did not influence when participants output the CIs.

Discussion

The results of Experiment 3 were largely concordant with Experiment 2. Most importantly, relational processing reduced false recall relative to reading. Although the generation task variants did not enhance correct recall overall, the recall dynamics analyses clarified that generation improved the recall of list items beyond the first few serial positions.

Cross-experiment analyses

Before turning to a more detailed consideration of our findings, we next report cross-experiment analyses of recall and recall dynamics.

Recall. To determine whether item-specific and relational task variants had similar effects across experiments, we first conducted separate 3(Group: Item-Specific vs. Relational vs. Read) × 3(Experiment: 1 vs. 2 vs. 3) between-group ANOVAs on correct recall, false recall of CIs, and extra-list intrusions. No significant interactions with experiment were found, largest $F(2, 312) = 1.70$, MSE = .01, $p = .15$. Importantly, the interaction effect for false recall was not reliable, $F < 1$, suggesting that the pattern of false recall in Experiment 1 (where relational processing did not reliably reduce the DRM illusion) was similar to the pattern in Experiments 2 and 3 (where the reductions were reliable).

Given these similarities across these three measures, we next conducted a series of analyses pooled across task variants to provide more powerful tests of potential differences between the effects of item-specific and relational processing on recall. Figure 1 (bottom panel) displays the pooled proportions of correct recall, false recall, and mean number of extra-list intrusions per list.

Correct recall of list items differed across groups, $F(2, 318)=15.37$, MSE=.01, $\eta_p^2=.09$. Relative to the read group, correct recall was enhanced following item-specific processing (.65 vs. .71), $t(212)=4.77$, SEM=.01, $d=.66$, and relational processing (.65 vs. .71), $t(211)=5.11$, SEM=.01, $d=.70$; this enhancement was equivalent for item-specific and relational processing, $t<1$.

False recall also differed across groups, $F(2, 318)=13.91$, MSE=.06, $\eta_p^2=.08$. Relative to the read group, false recall was reduced by item-specific processing (.39 vs. .21), $t(212)=5.32$, SEM=.03, $d=.73$, and, most importantly, it was also reduced by relational processing (.39 vs. .27), $t(211)=3.11$, SEM=.03, $d=.43$. The pooled analyses also revealed that false recall of CIs was more likely following relational processing than item-specific processing (.27 vs. .21). This difference was right at the conventional level of significance, $t(213)=1.95$, SEM=.03, $p=.05$, $d=.27$.

Mean number of extra-list intrusions per list also differed across groups, $F(2, 318)=4.50$, MSE=.08, $\eta_p^2=.03$. Relative to reading, intrusions were reduced following both item-specific processing (.27 vs. .15), $t(212)=2.65$, SEM=.03, $d=.36$, and relational processing (.27 vs. .18), $t(211)=2.17$, SEM=.03, $d=.30$. Importantly, intrusions were equivalent following item-specific and relational processing (.15 vs. .18), $t<1$, suggesting that both types of processing led to more stringent monitoring of recall output at test, reducing intrusion rates.

Recall dynamics. Next, we evaluated whether there were differences across the experiments in serial position recall using a 3(Group) × 3(Experiment) × 12(List Position) mixed ANOVA. This analysis revealed a significant three-way interaction, $F(44, 3443)=1.69$, MSE=.02, $\eta_p^2=.02$. As Figure 2 illustrates, this interaction reflected the different serial position curves in Experiment 3 relative to Experiments 1 and 2, as discussed above. We nonetheless pooled across experiments to highlight the cross-experiment consistencies apparent in the bottom panel of Figure 2. Primacy and recency effects occurred for all groups, and recall was higher for the item-specific and relational groups than for the read group across the middle serial positions. These effects were confirmed by main effects of group, $F(2, 318)=11.27$, MSE=.11, $\eta_p^2=.07$; serial position, $F(11, 3509)=479.96$, MSE=.01, $\eta_p^2=.58$; and a significant interaction, $F(22, 3509)=9.82$, MSE=.01, $\eta_p^2=.09$. Moreover, the read group showed a larger primacy effect for the first serial position, and the item-specific and relational groups showed greater recency effects for the last two serial positions based on Bonferroni-corrected *t* tests, all $ps<.03$.

The mean output position for CIs was also compared. Although group processing differences were not found in the individual experiments, the pooled analysis showed a marginal group effect, $F(2, 231)=2.51$, MSE=4.04, $p=.08$, $\eta_p^2=.02$. Post hoc comparisons revealed that the mean output position of the CIs was very similar following item-specific processing as in the read groups (6.80 vs. 6.82), $t<1$. However, the CIs were recalled later following relational processing than in the read group (7.41 vs. 6.82), $t(167)=1.94$, SEM=.21, $p=.05$, $d=.30$, and were recalled marginally later than following item-specific processing (7.41 vs. 6.80), $t(146)=1.78$, SEM=.24, $p=.08$, $d=.29$. Thus, an interesting finding was that relational processing tended to delay the output of the CIs, though we hesitate to make strong conclusions given the marginal omnibus.

A marginal effect of experiment was also found, $F(2, 231)=2.73$, MSE=4.04, $p=.07$, $\eta_p^2=.02$. The mean output position of the CIs was similar in Experiments 1 and 2 (6.93 vs. 7.38), $t(170)=1.41$, SEM=.22, $p=.16$, and in Experiments 1 and 3 (6.93 vs. 6.64), $t<1$. However, the CIs tended to be output later in Experiment 2 (pleasantness ratings) than in Experiment 3 (anagram generation) (7.38 vs. 6.64), $t(147)=2.32$, SEM=.22, $d=.38$. We do not place emphasis on this difference as it may depend on how pleasantness versus generation tasks are instantiated. The interaction between group and experiment was not reliable, $F<1$.

General discussion

The primary goal of our experiments was to compare the effects of item-specific and relational processing on free recall in the DRM paradigm. Across experiments, item-specific and relational processing were engendered via processing instructions (Experiment 1), variations in a pleasantness rating task (Experiment 2), or the presence or absence of an associated cue word to alter how anagrams were solved (Experiment 3). Overall, compared with simply reading DRM lists, the item-specific and relational variants both boosted correct recall, replicating Huff and Bodner's (2013) findings when recognition memory was tested following the same three encoding manipulations. However, whereas Huff and Bodner found that only the item-specific variants reduced the DRM illusion in recognition, here we found evidence that *both* variants typically reduced the illusion in recall. Overall, then, item-specific processing yields a mirror effect pattern in both tasks, whereas relational processing only does so in recall.

Relational processing reduces false memory in recall but not recognition

Our study sets forth the challenge of explaining this novel dissociation in the effects of relational processing across recall and recognition tasks. Of course, there are fundamental differences in how memory is accessed and acted upon in recall versus recognition. Recall places greater reliance upon recollective processes, due to the self-generative nature of this task. In contrast, although recollective processes are likely employed during old/new recognition, participants can also respond based on a test item's familiarity (Mandler, 1980; Yonelinas, 2002). The relational processing variants we use were designed to lead participants to generate associations among the words on a given list. This emphasis likely increased activation (Roediger et al., 2001) and/or led to the encoding of more gist-based information (Brainerd & Reyna, 2002) associated with the CIs during study relative to reading, as evidenced by the higher memory information index following relational processing found in recognition (see Huff et al., 2015, for a review and meta-analysis). The task dissociation we obtained is consistent with the possibility that relational encoding likely led the CIs to seem familiar, rather than recollected, when they came to mind during the recall test (cf. Brainerd, Payne, Wright, & Reyna, 2003, for evidence of recollection-based errors). If the relational groups output items on a recall task primarily based on recollection, then they might not output CIs that are merely experienced as familiar.

Of course, by this reasoning, the rate of false recall should have been lowest in the read groups—who should have been the least likely to experience the CIs as familiar (under the assumption that they did not perform as much relational processing at study). We propose that two factors likely contributed to the larger DRM illusion in the read groups than in the relational groups. First, the relational groups may have been more likely to experience recollection for the list items at test than the read groups, due to the elaborative encoding afforded by the relational processing. As a result, the relational groups would be less likely to have to resort to using "mere" familiarity as a basis for outputting items.

Indeed, there is evidence that the type of information participants use to justify recollection responses to one set of items can depend on the richness of their memory for another set of items. For instance, Bodner and Lindsay (2003) showed that whether a set of items studied in a medium level of processing task were deemed recollected versus familiar depended on whether they were mixed with deep level of processing items at test (akin to the relational groups in our studies) or with shallow level of processing items at test (akin to the read groups in our studies). Second, as shown in our pooled analyses, the extra-list intrusion rate was lower following relational processing than in the read condition, which is suggestive of greater memory monitoring in the relational group. Increased scrutiny of recall in the relational groups would also work to lower the DRM illusion relative to the read group and indeed may have prioritised recollection when retrieving items at test. Participants in the read group would thus be more likely to output CIs based on familiarity, granted they might also be less likely to experience them as familiar.

Thus, we suggest that participants may be more willing to provide an "old" judgement on a recognition test based on familiarity than they are to output an item based on familiarity during recall. Given our suggestion that relational processing boosts familiarity of the CIs, false recognition is higher following relational processing than it is in the read group. Importantly, this pattern holds despite elevated memory monitoring on recognition (e.g., Huff & Bodner, 2013). As Huff and Bodner discuss, memory monitoring at test is less successful following relational processing than after reading or item-specific processing because more information has been encoded about the CIs.

The line of reasoning presented here can also be extended to explain why the DRM illusion in recall was reduced even more by item-specific processing than by relational processing in the pooled analysis, even though the level of memory monitoring was similar. By extension, then, the reliance on familiarity during a recall test might be reduced even more following item-specific processing than following relational processing. Instead, item-specific processing should lead participants to rely on recollection at test (e.g., Yonelinas, 2002).

We acknowledge that our explanation of the task dissociation between recall and recognition is speculative given we do not have direct or indirect measures of familiarity/recollection and rests on several assumptions that future research should aim to test. First, it will need to be demonstrated that recall relies on recollection more than recognition, and that recognition relies on familiarity more than recall. This is a difficult proposition to evaluate, given the two tasks yield different dependent measures. One solution may be to collect remember/know judgements in both tasks—including for recalled items (e.g., see Mickes, Seale-Carlisle, & Wixted, 2013)—and to then compare effects between recall and recognition tasks. Second, the claim that relational processing increases the familiarity of CIs more than reading warrants testing. For example, relational processing may help participants detect CIs during encoding, allowing those items to be excluded when monitoring during recall. This possibility is consistent with the reduced false recall following relational processing relative to reading. To test this possibility, one could employ an inclusion recall to determine whether participants are consciously aware of the CI, but choosing not to report it on recall. On this test, participants would be asked to output all items that come to mind during recall (to minimise memory monitoring) so that familiar items would be output by all three groups. On this test, more CIs may be output following relational processing than in the read group—despite lower false recall levels in a standard recall task.

Our experiments revealed a consistent pattern in which both item-specific and relational processing tasks induced a mirror effect relative to a read control. In contrast, compared with a read condition, the item-specific tasks used in prior studies have sometimes reduced false recall (Burns et al., 2007; Thomas & Sommers, 2005) and sometimes inflated false recall (Butler et al., 2010), and they have not always improved correct recall (Butler et al., 2010; Thomas & Sommers, 2005). We suggest that some of these discrepancies may be due to task differences in how processing was recruited. For instance, embedding DRM words in sentences, as was done by Thomas and Sommers, may have drawn attention away from encoding the list items, thus reducing correct recall. Butler et al. reported that participants in their item-specific group frequently mentioned list word characteristics that were not distinctive, but were instead related to the CI. Thus, their "item-specific" instructions sometimes induced relational processing, a pattern consistent with the elevated false recall they reported. With our item-specific instructions (Experiment 1), the experimenter corrected participants who were generating words related to the CI during the speak-aloud practice list.

A related caveat is that although our item-specific and relational processing task variants were designed to encourage one type of processing over the other, we do not claim that either variant was process pure (see also Huff & Bodner, 2013). Indeed, lists that share a strong category structure such as DRM lists likely induce more relational processing than lists of unrelated items (Huff & Bodner, 2014; Hunt & Seta, 1984). Thus, our item-specific tasks must overcome participants' natural inclination towards relational processing before they can yield the expected memory benefits. Conversely, nothing prevents participants from engaging in some degree of item-specific processing as they perform our relational variants. This issue is particularly germane in Experiments 2 and 3, where we attempted to embed relational processing in traditional item-specific processing tasks. These relational processing variants needed to overcome any residual item-specific processing used to perform these tasks.

To date, item-specific and relational task variants have been compared only using tasks that are traditional item-specific tasks (e.g., pleasantness, generation). An important question for future research is whether the outcome patterns reported here for recall, and in Huff and Bodner (2013) for recognition, can also be obtained using variants of tasks that "normally" promote traditional relational processing tasks (e.g., associative judgements, category sorting). Moreover, it would be informative to examine other metrics of processing, such as category clustering, which may be indicative of the relative contributions of relational processing (Huff & Bodner, 2014; Hunt & Seta, 1984). Because individual DRM lists provide only a single semantic category, multiple lists would need to be blocked at study to allow clustering to be examined.

A case for measuring recall dynamics

Huff and Bodner (2013) estimated quantitative differences in the influence of item-specific and relational processing at encoding and retrieval using signal-detection indices. Our use of a free-recall task in the present study precluded this approach. The secondary goal of our study was therefore to explore several measures of recall dynamics to see whether item-specific and relational processing have similar effects on how DRM lists are recalled. Recall dynamics can provide insights into aspects of the recall process that are not revealed when examining recall proportions alone (e.g., Burns et al., 2007; Burns & Schoff, 1998; Kahana & Miller, 2013). Comparisons of serial position curves as a function of processing type performed at encoding are relatively rare (cf. Seiler & Engelkamp, 2003). Thus, our comparisons of these functions across item-specific, relational, and read-only groups in the DRM paradigm are novel.

For correct recall, item-specific and relational processing improved correct recall of DRM items from the middle list positions, relative to reading. This pattern held even when overall recall was equivalent across groups in Experiment 3. This pattern could reflect enhanced encoding of the list items and/or more focused retrieval efforts (i.e., memory monitoring) at test. Our findings cannot adjudicate between these options. Turning to false recall, the average output position of CIs was late in recall (replicating Roediger & McDermott, 1995), and novelly, the CI tended to be output later following relational processing relative to the other groups. We refrain from too much interpretation of this pattern pending replication, given the marginal main effect. However, the overall late output of CIs across conditions is likely related to participants exhausting available list items in memory and the exaggerated delayed output following relational processing may be indicative of qualitative differences in monitoring (vs. item-specific processing or reading). Given this potential effect, analyses of the output position of the CI may be useful in future work as an additional metric of how various manipulations affect CI false recall. Finally, though extra-list intrusions were rare, the pooled analysis revealed that intrusions were reduced following both item-specific and relational processing, relative to reading. This pattern provides some evidence of increased memory monitoring of candidate recall items at test, akin to what we find using our signal-detection approach in recognition. In this case, less effective diagnostic or disqualifying monitoring process (Gallo, 2004) would increase the likelihood of reporting intrusions. Thus, although the recall task prevented us from relying on our usual quantitative estimates of encoding and retrieval processes, other recall metrics can prove useful clues regarding how both processes shape correct and false recall.

Conclusion

We found that both item-specific and relational processing can improve memory accuracy relative to a read-control task when memory is tested via free recall. Our analyses of recall dynamics suggest that the increase in accuracy from these forms of processing is due in part to elevated monitoring and may reflect qualitative differences in how DRM lists are encoded. This pattern contrasts with the pattern in recognition in which a mirror effect pattern (increased correct memory and decreased false memory) has only been found following item-specific processing. We have suggested that this task dissociation may reflect differences across the two tasks in their relative emphasis on recollection (greater in recall) versus familiarity (greater in recognition). Further unpacking of the effects of these types of encoding across memory tasks will enhance our understanding of the factors that determine memory accuracy.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was supported by the Natural Sciences and Engineering Research Council of Canada through Discovery Grant RGPIN 238599-2015 to G.E.B.

Supplementary material

The Supplementary Material is available at qjep.sagepub.com.

References

Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language*, *48*, 1–15. doi:10.1016/S0749-596X(02)00518-1

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 941–947. doi:10.1037/0278-7393.27.4.941

Bodner, G. E., Huff, M. J., Lamontagne, R. W., & Azad, T. (2017). Getting at the source of distinctive encoding effects in the DRM paradigm: Evidence from signal-detection measures and source judgments. *Memory*, *25*, 647–655. doi:10.1080/09658211.2016.1205094

Bodner, G. E., & Lindsay, D. S. (2003). Remembering and knowing in context. *Journal of Memory and Language*, *48*, 563–580. doi:10.1016/S0749-596X(02)00502-8

Brainerd, C. J., Payne, D. G., Wright, R., & Reyna, V. F. (2003). Phantom recall. *Journal of Memory and Language*, *48*, 445–467.

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, *10*, 3–47. doi:10.1016/0273-2297(90)90003-M

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129. doi:10.1006/jecp.1998.2464

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*, 164–169. doi:10.1111/1467-8721.00192

Burns, D. J., Jenkins, C. L., & Dean, E. E. (2007). Falsely recalled items are rich in item-specific information. *Memory & Cognition*, *35*, 1630–1640. doi:10.3758/BF03193497

Burns, D. J., & Schoff, K. M. (1998). Slow and steady often ties the race: Effects of item-specific and relational processing on cumulative recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1041–1051. doi:10.1037/0278-7393.24.4.1041

Butler, K. M., McDaniel, M. A., McCabe, D. P., & Dornburg, C. C. (2010). The influence of distinctive processing manipulations on older adults' false memory. *Aging, Neuropsychology, and Cognition*, *17*, 129–159. doi:10.1080/13825580903029715

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428. doi:10.1037/0033-295X.82.6.407

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

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22. doi:10.1037/h0046671

Foley, M. A., Wozniak, K. H., & Gillum, A. (2006). Imagination and false memory inductions: Investigating the role of process, content, and source of imaginations. *Applied Cognitive Psychology*, *20*, 1119–1141. doi:10.1002/acp.1265

Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 120–128. doi:10.1037/0278-7393.30.1.120

Gallo, D. A. (2006). *Associative memory illusions*. New York, NY: Psychology Press.

Gallo, D. A., Roediger, I. I. I., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, *8*, 579–586. doi:10.3758/BF03196194

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 5–16. doi:10.1037/0278-7393.16.1.5

Gunter, R. W., Bodner, G. E., & Azad, T. (2007). Generation and mnemonic encoding induce a mirror effect in the DRM paradigm. *Memory & Cognition*, *35*, 1083–1092. doi:10.3758/BF03193480

Hege, A. C. G., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both impoverished relational-encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 787–795. doi:10.1037/0278-7393.30.4.787

Hicks, J. L., & Marsh, R. L. (2001). False recognition occurs more frequently during source identification than during old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 375–383. doi:10.1037/0278-7393.27.2.375

Hockley, W. E., & Cristi, C. (1996). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition*, *24*, 202–216. doi:10.3758/BF03200881

Huff, M. J., & Bodner, G. E. (2013). When does memory monitoring succeed versus fail? Comparing item-specific and relational encoding in the DRM paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1246–1256. doi:10.1037/a0031338

Huff, M. J., & Bodner, G. E. (2014). All varieties of encoding variability are not created equal: Separating variable processing from variable tasks. *Journal of Memory and Language*, *73*, 43–58. doi:10.1016/j.jml.2014.02.004

Huff, M. J., Bodner, G. E., & Fawcett, J. M. (2015). Effects of distinctive encoding on correct and false memory: A meta-analytic review of costs and benefits and their origins in the DRM paradigm. *Psychonomic Bulletin & Review*, *22*, 349–365. doi:10.3758/s13423-014-0648-8

Huff, M. J., Coane, J. H., Hutchison, K. A., Grasser, E. B., & Blais, J. E. (2012). Interpolated task effects on direct and mediated false recognition: Effects of initial recall, recognition, and the ironic effect of guessing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1720–1730. doi:10.1037/a0028476

Huff, M. J., McNabb, J., & Hutchison, K. A. (2015). List blocking and longer retention intervals reveal an influence of gist processing for lexically ambiguous critical lures. *Memory & Cognition*, *43*, 1193–1207. doi:10.3758/s13421-015-0533-3

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3–25). New York, NY: Oxford University Press.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 497–514. doi:10.1016/S0022-5371(81)90138-9

Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General*, *109*, 49–74. doi:10.1037/0096-3445.109.1.49

Hunt, R. R., & Seta, C. E. (1984). Category size effects in recall: The roles of relational and individual item information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 454–464. doi:10.1037/0278-7393.10.3.454

Hunt, R. R., & Smith, R. E. (2014). How distinctive processing enhances hits and reduces false alarms. *Journal of Memory and Language*, *75*, 45–57. doi:10.1016/j.jml.2014.04.007

Hunt, R. R., Smith, R. E., & Dunlap, K. R. (2011). How does distinctive processing reduce false recall? *Journal of Memory and Language*, *65*, 378–389. doi:10.1016/j.jml.2011.06.003

Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory and Language*, *52*, 1–28. doi:10.1016/j.jml.2004.08.003

Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review*, *4*, 577–581. doi:10.3758/BF03214352

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28. doi:10.1037/0033-2909.114.1.3

Kahana, M. J., & Miller, J. F. (2013). Memory recall, dynamics. In H. Pashler (Ed.), *Encyclopedia of the mind* (*Vol. 2*, pp. 493–497). Thousand Oaks, CA: SAGE.

Kelley, M. R., & Nairne, J. S. (2001). von Restorff revisited: Isolation, generation, and memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 54–66. doi:10.1037.0278-7393.27.1.54

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–514. doi:10.1037/0033-295X.103.3.490

Lampinen, J. M., Neuschatz, J. S., & Payne, D. G. (1999). Source attributions and false memories: A test of the demand characteristics account. *Psychonomic Bulletin & Review*, *6*, 130–135. doi:10.3758/BF03210820

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, *96*, 25–57.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271. doi:10.1037/0033-295X.87.3.252

McCabe, D. P., & Smith, A. D. (2002). The effect of warnings on false memories in young and older adults. *Memory & Cognition, 30*, 1065-1077. doi: 10.3758/BF03194324

McCabe, D. P., Presmanes, A. G., Robertson, C. L., & Smith, A. D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review*, *11*, 1074–1079. doi:10.3758/BF03196739

McCabe, D. P., & Smith, A. P. (2006). The distinctiveness heuristic in false recognition and false recall. *Memory*, *14*, 570–583. doi:10.1080/09658210600624564

McDaniel, M. A., Cahill, M. J., & Bugg, J. M. (2016). The curious case of orthographic distinctiveness: Disruption of categorical processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 104–113. doi:10.1037/xlm0000160

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212–230. doi:10.1006/jmla.1996.0012

Mickes, L., Seale-Carlisle, T. M., & Wixted, J. T. (2013). Rethinking familiarity: Remember/know judgments in free recall. *Journal of Memory and Language*, *68*, 333–319. doi:10.1016/j.jml.2013.01.001

Multhaup, K. S., & Conner, C. A. (2002). The effects of considering non-list sources on the Deese-Roediger-McDermott memory illusion. *Journal of Memory and Language*, *47*, 214–228. doi:10.1016/S0749-596X(02)00007-4

Nelson, D. L., McEvoy, C. L., & Schreiber, T. (1999). *The University of South Florida free association norms*. Retrieved from http://w3.usf.edu/FreeAssociation

Neuschatz, J. S., Payne, D. G., Lampinen, J. M., & Toglia, M. P. (2001). Assessing the effectiveness of warnings and phenomenological characteristics of false memories. *Memory*, *9*, 53–71. doi:10.1080/09658210042000076

Oliver, M. C., Bays, R. B., & Zabrucky, K. M. (2016). False memories and the DRM paradigm: Effects of imagery, list, and test type. *The Journal of General Psychology*, *143*, 33–48. doi:10.1080/00221309.2015.1110558

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, *35*, 261–285. doi:10.1006/jmla.1996.0015

Robin, F. (2010). Imagery and memory illusions. *Phenomenology and the Cognitive Sciences*, *9*, 253–262. doi:10.1007/s11097-010-9159-x

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

Roediger, H. L., III, & Guynn, M. (1996). Retrieval processes. In E. L. Bjork & R. A. Bjork (Eds.), *Human memory* (pp. 197–236). San Diego, CA: Academic Press.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814. doi:10.1037/0278-7393.21.4.803

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *40*, 1–24. doi:10.1006/jmla.1998.2611

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19*, 523–542. doi:10.3758/BF03197149

Seiler, K. H., & Engelkamp, J. (2003). The role of item-specific information for the serial position curve in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 954–964. doi:10.1037/0278-7393.29.5.954

Smith, R. E., Hunt, R. R., & Dunlap, K. R. (2015). Why do pictures, but not visual words, reduce older adults' false memories? *Psychology and Aging*, *30*, 647–655. doi:10.1037/pag0000044

Thomas, A. K., & Sommers, M. S. (2005). Attention to item-specific processing eliminates age effects in false memories. *Journal of Memory and Language*, *52*, 71–86. doi:10.1016/j.jml.2004.08.001

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

von Restorff, H. (1933). Uber die wirkung von bereichsbildungen im spurenfeld [On the effect of spheres formations in the trace field]. *Psychologiche Forschung*, *18*, 299–342.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517. doi:10.1006/jmla.2002.2864