Memory & Cognition (2018) 46:1287–1301
https://doi.org/10.3758/s13421-018-0837-1

Item-specific processing reduces false recognition in older and younger adults: Separating encoding and retrieval using signal detection and the diffusion model

Mark J. Huff[1] · Andrew J. Aschenbrenner[2]

Published online: 29 June 2018

Abstract
Our study examined processing effects in improving memory accuracy in older and younger adults. Specifically, we evaluated the effectiveness of item-specific and relational processing instructions relative to a read-only control task on correct and false recognition in younger and older adults using a categorized-list paradigm. In both age groups, item-specific and relational processing improved correct recognition versus a read-only control task, and item-specific encoding decreased false recognition relative to both the relational and read-only groups. This pattern was found in older adults despite overall elevated rates of false recognition. We then applied signal-detection and diffusion-modeling analyses, which separately utilized recognition responses and the latencies to those responses to estimate contributions of encoding and monitoring processes on recognition decisions. Converging evidence from both analyses demonstrated that item-specific processing benefits to memory accuracy were due to improvements of both encoding (estimates of *d*′ and drift rate) and monitoring (estimates of lambda and boundary separation) processes, and, importantly, occurred similarly in both younger and older adults. Thus, older and younger adults showed similar encoding-based and test-based benefits of item-specific processing to enhance memory accuracy.

Keywords Item-specific processing · Relational processing · Distinctiveness · Signal detection · Diffusion modeling

Processing information based on distinctive features often yields a memorial advantage. Distinctive processing refers to the processing of "difference within a context of similarity" (Hunt, 2006, p. 12; see, too, Schmidt, 1991, for discussion). A similar context occurs when a set of to-be-remembered materials contain shared spatial, temporal, or semantic features. Distinctive processing can therefore occur in a variety of contexts and study tasks. Benefits have been found perceptually, as evidenced by improved memory for information studied in picture versus word formatting (Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999), for words that are orthographically bizarre versus typical (Hunt & Elliot, 1980; McDaniel, Cahill, & Bugg, 2016), and for items studied in isolation (Kelley & Nairne, 2001; von Restorff, 1933). By extension, distinctive effects have also been found when study tasks encourage the encoding of unique features. Item-specific processing, for instance, has facilitated correct memory performance while simultaneously reducing false recognition to associatively related lure items, producing a net benefit to overall memory accuracy (Huff & Bodner, 2013; Hunt, Smith, & Dunlap, 2011). The purpose of the current study is to further assess the locus of accuracy-enhancing benefits due to item-specific processing by evaluating the contributions of study-based and test-based processes through signal-detection and diffusion-modeling analyses, and to determine whether these mechanisms differ between younger and older adults.

A common paradigm for examining false memory errors is the Deese–Roediger–McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) paradigm. The DRM paradigm presents participants with lists of associates at study (e.g., *bed, tired, rest*) that converge upon a single nonstudied critical lure (e.g., *sleep*). At test, false memory for lures is

Electronic supplementary material The online version of this article (https://doi.org/10.3758/s13421-018-0837-1) contains supplementary material, which is available to authorized users.

✉ Mark J. Huff
mark.huff@usm.edu

[1] Department of Psychology, The University of Southern Mississippi, Hattiesburg, MS 39406, USA

[2] Department of Neurology, Washington University in St. Louis, St. Louis, MO, USA

high: False recall approaches 50% (Roediger & McDermott, 1995) and false recognition can meet or exceed hit rates (Gallo, 2006, 2010 for review, Lampinen, Neuschatz, & Payne, 1999). Effects similar to the DRM illusion have been found using other associative materials. Studying categorically related words (e.g., *birds*) also results in false recall and false recognition for strongly related members (e.g., *robin, cardinal*; Huff & Bodner, 2014; Meade & Roediger, 2006).

Given the magnitude of false memory effects, it is important to identify reliable methods by which rates of false memory can be reduced. Successful methods include repeated study presentations (Benjamin, 2001; McDermott, 1996), and warnings presented before study and/or test (Gallo, Roediger, & McDermott, 2001; McCabe & Smith, 2002; Neuschatz, Payne, Lampinen, & Toglia, 2001). Importantly, distinctive encoding has also been successful at reducing the illusion. Israel and Schacter (1997) reported that when studying DRM lists that are accompanied by a picture of the words referent, the DRM illusion is reduced relative to a condition in which the word is presented in isolation. This pattern has similarly been found using other distinctive manipulations, such as studying words in unique fonts (Arndt & Reder, 2003), generating words from anagrams (Gunter, Bodner, & Azad, 2007; McCabe & Smith, 2006), creating mental images of list words (Foley, Wozniak, & Gillum, 2006; Oliver, Bays, & Zabrucky, 2016; Robin, 2010), and, relevant to the present study, processing the unique or distinctive features of individual study words using item-specific processing study tasks (Huff & Bodner, 2013; McCabe, Presmanes, Robertson, & Smith, 2004; Smith & Hunt, 1998). These benefits are particularly noteworthy because they often induce a *mirror effect* (Glanzer & Adams, 1990)—an increase in correct memory coupled with a decrease in false memory for distinctive encoding relative to a read-only or processing-neutral control task.

Separating encoding and retrieval processes in distinctive encoding

The processes that give rise to a distinctive encoding advantage in memory are not fully understood. At least two mechanisms have been proposed to account for the benefits of distinctive processing: Those that facilitate distinctive encoding through processes at study, and those that enhance strategic monitoring at test. According to the *impoverished relational encoding* account (Hege & Dodson, 2004; Hockley & Cristi, 1996), distinctive processing operates at study by either disrupting the thematic consistency of the list (Brainerd & Reyna, 2002), or by reducing the implicit associations shared between list items and the critical lure (Roediger, Balota, & Watson, 2001). However, distinctive processing may influence retrieval by leading participants to adopt a global-monitoring strategy known as the *distinctiveness heuristic*, in which participants apply a test-based decision rule in which memory items are reported only when accompanied by the recollection of distinctive details. Distinctive details that are recollected can provide diagnostic evidence that the item was studied, whereas absence of these details can provide diagnostic evidence that an item was not studied (see Gallo, 2004, 2010, for discussion).

One approach for separating these processes is using a within-subjects design, in which participants study and are tested on two types of DRM lists: Those studied in a distinctive format and those that are not. The impoverished relational encoding hypothesis would predict a selective reduction in false memory only for the distinctive lists, whereas the global distinctiveness heuristic would equally influence both lists. Unfortunately, results from such studies have not yielded consistent results. Schacter et al. (1999) had participants study DRM lists in which half included pictures presented alongside words, whereas the other half consisted of words in isolation. At test, the DRM illusion was equally low for lists studied under distinctive and nondistinctive formats, consistent with use of a global distinctiveness heuristic. In contrast, Arndt and Reder (2003) showed a reduction in the DRM illusion selectively for lists studied using distinctive fonts, a pattern consistent with that of impoverished relational encoding.

Another approach to separating these mechanisms is through test-based inclusion instructions. Inclusion instructions require participants to report or endorse all test items that were studied or are related to items that were studied (Gunter et al., 2007; Hege & Dodson, 2004; Hunt et al., 2011). These instructions should therefore reduce the contributions of test-based monitoring (i.e., the distinctiveness heuristic), leaving only encoding-based processes in place. Again, however, inclusion instructions have yielded reductions in the DRM illusion selectively for distinctive lists (Hege & Dodson, 2004; Hunt et al., 2011), consistent with an encoding locus, but have also produced a null effect, consistent with test-based monitoring (Gunter et al., 2007; Pierce, Gallo, Weiss, & Schacter, 2005).

Aside from inconsistent effects, these approaches also suffer an additional interpretation issue. Specifically, evidence found for either the distinctiveness heuristic or impoverished relational encoding eliminates the other process by default, which excludes a possibility that *both* encoding-based and test-based processes may be involved in the distinctiveness reduction. More recently, it has been argued that a signal-detection approach applied to recognition is a more appropriate method for discerning the contributions of encoding and retrieval processes in distinctive processing (Gunter et al., 2007; Huff & Bodner, 2013; Huff, Bodner, & Fawcett, 2015). Using this approach, signal detection allows for a separation of underlying memory experiences for studied versus nonstudied information (or discriminability, d') from memory monitoring, or the likelihood that nonstudied information is reported as studied versus correctly rejected. When applied to encoding-based and test-based processes in distinctive processing, discriminability estimates the amount of encoded memory information for studied list items or the critical lure (an index of impoverished relational encoding), whereas an estimate of memory monitoring based solely on the false alarm rate (computed as lambda, λ), quantitatively estimates the amount of monitoring used at test (an index of the distinctiveness heuristic; see Fig. 1, top panel, for a graphical depiction of *d*′ and lambda in the distributions). Thus, signal-detection estimates provide separate indices of encoding and retrieval, allowing for the possibility that both processes may contribute to distinctiveness effects.[1]

Huff and Bodner (2013) utilized a signal-detection approach to estimate encoding and retrieval contributions in correct and false recognition in three experiments using a between-subjects design. In each experiment, two variants of a deep-processing task were compared with a read-only control group. One variant encouraged item-specific processing, or the processing of unique features of each DRM list item, whereas the other encouraged relational, or shared characteristics of the list items (Hunt & Einstein, 1981). Item-specific and relational processing variants were created in three experiments, which used a processing instruction study task, a pleasantness-rating task, and an anagram-generation task. Across experiments, item-specific and relational variants consistently improved correct recognition over the read groups, and these improvements were due to a combination of increased encoding of studied list items and increased monitoring (elevated *d*′ and lambda estimates). False recognition following item-specific processing was lower than both the read and relational processing groups, and this distinctive reduction was due to a combination of *both* a reduction in encoded memory information for the critical lure and an increase in test-based monitoring (reduced *d*′ and elevated lambda estimates). Impoverished relational encoding and the distinctiveness heuristic therefore operated in concert as a two-stage process: Impoverished relational encoding disrupted initial critical-lure activation, and the distinctiveness heuristic enhanced diagnostic monitoring at test (see Hunt & Smith, 2014, for a similar proposed process). This signal-detection pattern was further supported in a meta-analysis that included data sets from early studies using distinctive study tasks reviewed above (Huff et al., 2015).

While signal-detection indices ostensibly provide separate measures of encoding and retrieval processes, it is important to note that these indices are only estimates based on hits and false alarms. A goal of our study was to test for convergent validity of the signal-detection measures by applying a computational model of binary choice, the drift diffusion model (Ratcliff, 1978). The diffusion model utilizes all aspects of performance, proportion of responses and reaction-time distributions, to derive latent, psychologically interpretable parameters. As shown in Fig. 1 (bottom panel), the diffusion model assumes the noisy accumulation of evidence toward a given response boundary (e.g., "old" or "new" in a standard episodic recognition paradigm). The model produces four primary parameters. The drift rate reflects the rate at which memory evidence accumulates toward a given response, which may reflect the amount or strength of information stored in memory. Boundary separation indicates how much evidence is required before a response is made. This parameter is typically interpreted as reflecting response caution. Nondecision time reflects processes that occur outside the decision process, such as stimulus decoding and response execution. Finally, the starting point reflects where information begins to accumulate and reflects preference for one point over another.

The drift rate, nondecision time, and start point all have associated variability parameters that are also estimated from the model. Although several parameters are derived, we are primarily interested in only two of them for the present purposes. Specifically, we hypothesized that the drift rate and *d*′ (discriminability) would similarly be affected by the type of processing task completed, producing an estimate of encoded memory information. Likewise, we hypothesized that boundary separation and lambda would also be similarly affected by processing and estimate memory monitoring. In addition, the ability of the diffusion model to provide novel insights into age-related differences has been well established (Ratcliff, Thapar, & McKoon, 2004, 2010). Furthermore, it is often the case that diffusion-model parameters are more sensitive to underlying changes than are summary statistics such as overall accuracy (Aschenbrenner, Balota, Gordon, Ratcliff, & Morris, 2016; White, Ratcliff, Vasey, & McKoon, 2010). These parameters may also be sensitive to item-specific encoding processes in older adults, a topic we will now discuss.

Item-specific processing effects in aging

An additional goal of our study was to evaluate the effects of item-specific and relational processing relative to reading in

[1] Previous research using the DRM paradigm has examined the role of general criterion shifts as a causal mechanism in false recognition (Miller & Wolford, 1999; Wixted & Stretch, 2000). In these papers, criterion is computed using a bias measure (e.g., *c*), which, for false recognition, captures the propensity to respond "old" versus "new" to critical lures and is computed as the intersection between the critical item and critical item control distributions. We (and others, Gunter et al., 2007; Huff et al., 2015) have suggested that traditional response bias measures may be less accurate measures of monitoring, because the distribution of old responses to critical items can shift due to changes in associative or thematic activation of the critical item (which would shift the "hit" distribution), due to changes in memory monitoring, or both. Thus, bias measures in signal detection that utilize both distributions are ambiguous regarding the underlying cause. In contrast, lambda is computed by using only the false-alarm rate to controls and mathematically is less affected by a shifting "hit" distribution for critical items. We therefore argue that lambda is a better estimate of memory monitoring over traditional bias measures, given hits are not mathematically factored into computing the estimate. Of course, we do not argue that lambda is a "pure" measure of monitoring and note that monitoring and bias are likely correlated. Separation of the two via quantitative recognition responses is difficult and may require more qualitative memorial information such as metamemory judgments.

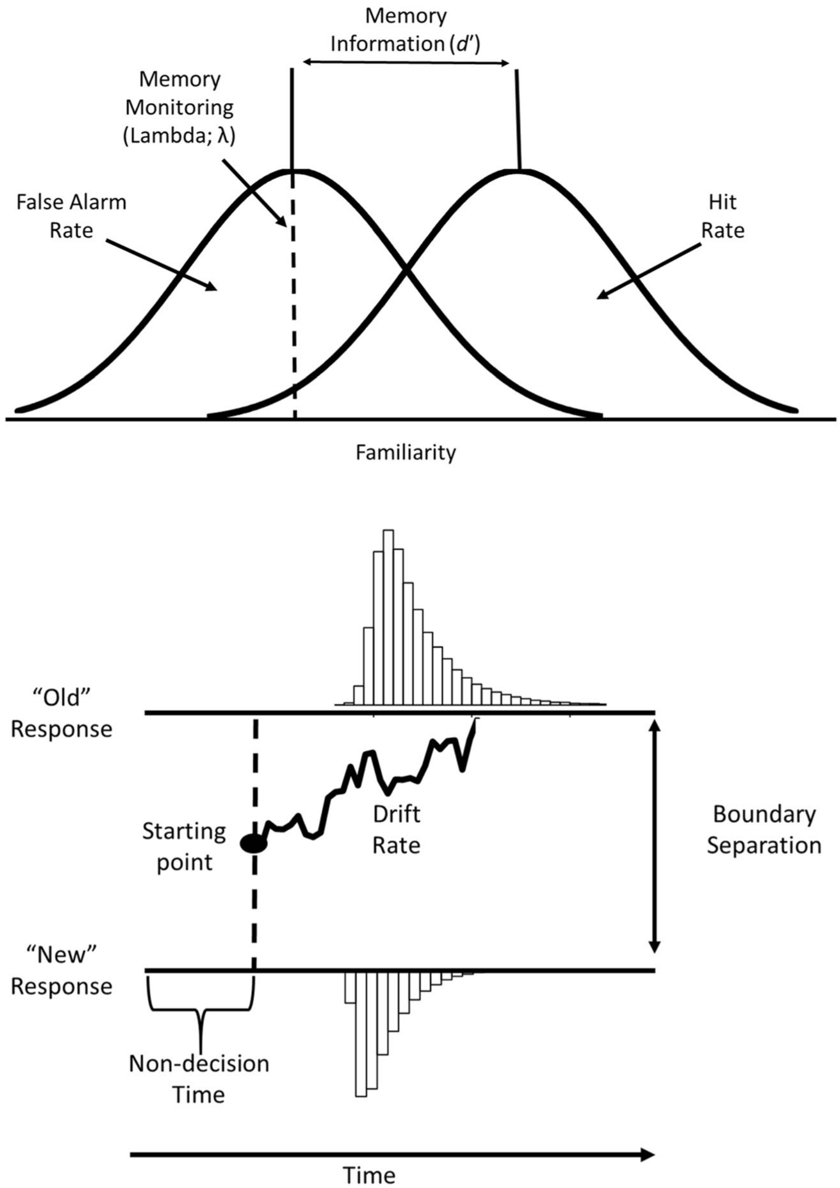

Fig. 1 Parameters of the signal-detection model (top panel) and the diffusion model (bottom panel) and their relations to hit and false-alarm rates and to old and new recognition responses, respectively. For the signal-detection model, the hit distribution applies to correctly studied list items and critical lures, whereas the false-alarm distribution applies to list-item controls and critical-lure controls

both younger and older adults.[2] Older adults often show episodic memory deficits (see Balota, Dolan, & Duchek, 2000; Kausler, 1994, for reviews), which are generally larger on tests of free recall that require self-initiated retrieval than recognition (Craik & McDowd, 1987; Wahlheim & Huff, 2015). In the DRM paradigm, older adults have also been shown to be more susceptible to the DRM illusion (Balota et al., 1999; Norman & Schacter, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998), demonstrating that age-related memory declines compromise overall memory accuracy. Given the benefits of item-specific processing in younger adults, a critical question is whether item-specific processing effects extend to older adults to remedy declines in memory accuracy.

Previous research examining item-specific effects in older adults in the DRM paradigm have been relatively mixed, though many of these studies have encouraged item-specific processing by manipulating perceptual features of DRM list items. As reviewed by Smith, Hunt, and Dunlap (2015), some studies have shown a reduction in older adults when DRM

[2] We note that other studies (e.g., Hunt et al., 2011, Experiment 1; Toglia, Neuschatz, & Goodwin, 1999; Thapar & McDermott, 2001) have evaluated item-specific processing (e.g., pleasantness ratings) in the context of a shallow levels-of-processing task (e.g., vowel counting). These studies have shown that item-specific tasks increase both correct and false recognition relative to the shallow levels-of-processing task. Given the surface-level processing conducted with shallow tasks, we argue that the read-only task is a more appropriate processing-neutral control task. As Hunt et al. (2011) demonstrate, the comparison group is crucial for demonstrating whether or not item-specific processing will show a reduction in false memory.

words are studied in picture format versus words (Schacter et al., 1999), while others have found no reduction (Gallo, Bell, Beier, & Schacter, 2006), in both recall and recognition. Other studies however, have encouraged item-specific processing at the task level. Thomas and Sommers (2005) had older adults study DRM list words by embedding them in sentences that either converged (i.e., relational) or diverged (i.e., item-specific) with a list theme relative to a read-only control condition. Divergent sentences were found to reduce false recognition, a pattern consistent with item-specific effects in younger adults. More relevant to the present study, Butler, McDaniel, McCabe, and Dornburg (2010) had older adults either verbally generate unique characteristics of DRM words or read DRM words aloud. In contrast to item-specific effects found in younger adults (Gunter et al., 2007; Huff & Bodner, 2013; McCabe et al., 2004), DRM false recall was greater following item-specific instructions than reading. Further, most studies above failed to find correct memory benefits following item-specific processing, suggesting that item-specific benefits may be diminished in older (vs. younger) adults (see Smith, 2006, for a similar conclusion). The lack of item-specific benefits may therefore reflect deficits of item-specific processing to engage encoding processes, monitoring processes, or both, processes that have not yet been quantified in this context.

The current study

Given the mixed findings regarding item-specific processing in older adults, we sought to further evaluate the effects of item-specific processing relative to read-control and relational processing tasks in both younger and older adults (cf. Butler et al., 2010). We utilized signal-detection and diffusion modeling approaches to determine whether changes in correct and false recognition following item-specific and relational processing reflect effects of encoding, retrieval, or a combination of the two. To obtain stable estimates of diffusion-model parameters for critical lures, we utilized categorized (rather than DRM) study lists, which allow for an increased number of critical lures per list by withholding the top most common exemplars from study. Categorized list words share strong thematic and associative relations and therefore operate similarly to induce false recognition as DRM lists.

In the present work, younger and older adults studied a series of categorized word lists followed immediately by an old/new recognition test. Participants studied lists using either item-specific instructions or relational instructions which were then compared to a read-only control group. We expected that item-specific and relational processing instructions would enhance correct recognition relative to reading given both tasks involve a semantically deep levels-of-processing. Based on Huff and Bodner's (2013) results, we expected that item-specific processing would reduce false recognition relative to relational processing and reading. Huff and Bodner did not find an increase in false recognition over reading, which was suggested to occur due to relational processing increasing the detection and rejection of DRM critical lures at study and test. Given our use of categorized lists, the lures would be less likely to be detected. Therefore, we predicted that relational processing would increase false recognition over reading despite an improvement in correct recognition.

Additionally, we predicted that older (vs. younger) adults would show limited accuracy gains from item-specific processing given the extant literature (Butler et al., 2010; Smith, 2006); however, a counter prediction could also be made that older adults may instead show benefits of item-specific processing over reading. This latter prediction stems from the findings that older adults show deficits at spontaneously engaging deep encoding processes given limited attentional resources (e.g., Anderson, Craik, & Naveh-Benjamin, 1998; Craik & Byrd, 1982), but when given an encoding strategy accompanied by practice and feedback, they may show benefits. Finally, signal detection and diffusion modeling were expected to produce converging estimates of encoding and monitoring processes. We expected that the mirror effect found following item-specific processing would reflect a combination of enhanced encoding processes (increased encoded information for list items; decreased encoded information for critical lures) and an increase in test-based monitoring as estimated through signal-detection and diffusion-modeling analyses.

Method

Participants

Eighty-eight younger adults and 82 older adults were recruited for participation. Data from three younger and four older adults were eliminated due to failures to follow instructions either during the study or recognition phases, leaving 85 younger and 78 older adults for analysis. Younger adults were psychology undergraduates who completed the experiment for course credit ($M_{\text{age}} = 21.00$ years, range: 18–39; 77.64% female). Older adults were recruited from either The University of Southern Mississippi's Osher Lifelong Learning Institute or from the greater Hattiesburg, MS, community ($M_{\text{age}} = 74.58$ years, range: 60–90; 67.95% female) and were compensated $10 per hour for their participation. Participants were randomly assigned to either the item-specific, relational, or read encoding groups, with distribution across groups being relatively even.

Demographics for both age groups are reported in Table 1. Older adults reported more years of education than younger

Table 1 Participant characteristics for younger and older adults

Variable/Group	Younger adults	Older adults
N	85	78
Age (years)	21.00 (2.17)	74.58 (6.77)
Education (years)	13.52 (2.17)	15.55 (2.76)
Shipley Vocabulary	26.53 (4.47)	32.27 (5.22)
Mini Mental State Exam	28.13 (1.75)	27.14 (2.06)

Note. Mean (*SD*)

adults (15.55 vs. 13.52), $t(159) = 5.32$, $SEM = .35$, and had higher vocabulary scores (32.27 vs. 26.53), $t(152) = 7.32$. $SEM = .78$, on the Shipley Institute of Living Scale (Shipley, 1986), but scored lower on the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) than younger adults (27.14 vs. 28.13), $t(152) = 3.21$, $SEM = .31$, though still within a normal range (24 or greater) based on standard scoring guidelines. Education information was not reported by one younger and one older adult, and the Shipley scale and MMSE were not completed by nine younger adults. Further, due to a computer error, response latencies were unavailable for four younger adults. These individuals were included in all analyses, aside from the diffusion model, as this analysis required latencies.

Materials

Study and test materials were presented via a computer running SuperLab software. Two sets of 10 categorized lists were constructed that contained items taken from either the Battig and Montague (1969) or Van Overschelde, Rawson, and Dunlosky (2004) categorical word norms. The top 20 exemplars from each category were used. Of these exemplars, the five most common exemplars (i.e., the top five items in the norms) were designated as critical lures and were not studied. The remaining 15 exemplars were then used as study items and presented in descending order of typicality (see Huff, Balota, & Hutchison, 2016; Meade & Roediger, 2006, for a similar procedure). A 260-item recognition test was then constructed and composed of 80 studied list items (from positions 1, 3, 5, 7, 9, 11, 13, and 15 from each studied list), 80 list item controls taken from the nonstudied list set (from the same list positions), 50 critical lures (five per studied list), and 50 critical-lure controls (five per nonstudied list). The recognition test was presented in a newly randomized order for each participant.

Procedure

All participants were tested individually with an experimenter present. Participants were instructed that they would study a series of word lists and that each word would be read aloud. The item-specific group was further instructed to "think of a unique characteristic that differentiates each word from the other words on the list." The relational group was instructed to "concentrate on what the words have in common to associate them together." Item-specific and relational instructions were modeled after those used by Huff and Bodner (2013). The read group simply read each word aloud. An eight-item practice list made of weak associates was then presented to participants (a mediated list taken from Huff & Hutchison, 2011). For each item on the practice list, participants were instructed to verbally state the unique or shared characteristic for each of the items. An experimenter provided participants with feedback to ensure the appropriate encoding task was deployed (cf. Butler et al., 2010). Participants then studied the 10 experimental lists, which were once-randomized and presented in the same order. Study duration for each list item was participant paced. Each list was separated by a screen with the words "next list," which the experimenter read aloud. The recognition test immediately followed the study phase. Words were displayed individually on the computer screen and participants classified each word as "old" if the word was presented on one of the previous study lists, or "new" if the word was not using a button-labeled response box. Given the importance of response latencies in our analyses, participants were instructed to place their index fingers directly on the old/new buttons and respond quickly to each test word, but without compromising accuracy. The recognition test was similarly participant paced.

Immediately following the recognition test, participants completed the demographics, vocabulary, and MMSE questionnaires. Participants were then debriefed and compensated with course credit or payment. An entire session lasted approximately 60 minutes.

Results

We set a $p < .05$ significance level for analyses, and provide effect sizes for reliable and marginal analyses of variance (ANOVA) effects (η_p^2) and *t* tests (Cohen's *d*). Mean recognition scores, signal-detection indices, and diffusion-model parameters for correctly recognized list items and falsely recognized critical lures are plotted in the referenced figures below, and numeric means are reported in the Appendix Table 2 for completeness. A post hoc power analysis using G*Power (Erdfelder, Faul, & Buchner, 1996) revealed that our sample size had high statistical power (.90) to detect medium-sized effects ($d = .40$).

Preliminary data analyses

Signal detection Recognition proportions of old responses were analyzed across list items, critical items, and their respective controls. A signal-detection measure of sensitivity was then applied to compute d' and lambda for list items and critical items. The d' index computed separately for list and critical items by taking the z score of the hit rate for each item type (with critical items treated as hits) minus the z score of the false-alarm rate for the appropriate control items for each participant. The lambda index was computed by taking the z score of 1 minus false-alarm rate for list item and critical item controls for each participant. For lambda, higher values are indicative of reduced false-alarm rates and greater monitoring (see Fig. 1, top panel, for a signal-detection distribution with parameters). Hit and false-alarm rates of 1.0 or 0 were adjusted using Macmillan and Creelman's (1991) 1/2n correction.

Diffusion-model fitting The diffusion model was fit to individual participant's data using Fast-DM Version 30.2 (Voss & Voss, 2007). Responses were trimmed by removing any that were faster than 200 milliseconds or slower than 8,000 milliseconds. Responses were then trimmed further by removing those that were greater than three standard deviations from the participant's mean response latency. This trimming procedure removed an average of 3% of trials across item types. The Kolmogorov–Smirnov (K–S) method was chosen as the optimization criterion. Briefly, the program recovers the optimal model parameters by minimizing the maximum distance between the empirical cumulative density function and the one predicted by the model. The upper response boundary was arbitrarily designated "old" and the bottom boundary as "new" (see Fig. 1, bottom panel). Drift rate and boundary separation were allowed to vary across item types. In the interest of parsimony, the other parameters were held constant across conditions. Interpretation of the parameters depends heavily upon how well the model describes the data. Fast-DM provides an estimate of individual model fit using the K–S statistic. With the exception of a single older adult, model fit was acceptable (K–S; $p > .05$). In order to maintain consistency with the signal detection analyses, the poor fitting participant was retained. Mean and median response latencies as a function of encoding task and age group as well as plots of model fits are available in the Supplemental Materials.

Correct recognition

Proportions of correct recognition as a function of encoding task and age group are reported in Fig. 2 (top panel). A 3(encoding: item-specific vs. relational vs. read) × 2(age: younger vs. older) between-subjects ANOVA was used to analyze the proportions of old responses to studied list items. An effect of encoding, $F(2, 157) = 23.52$, mean square error (MSE) = .02, $\eta_p^2 = .23$, revealed that correct recognition was greater in the item-specific group than

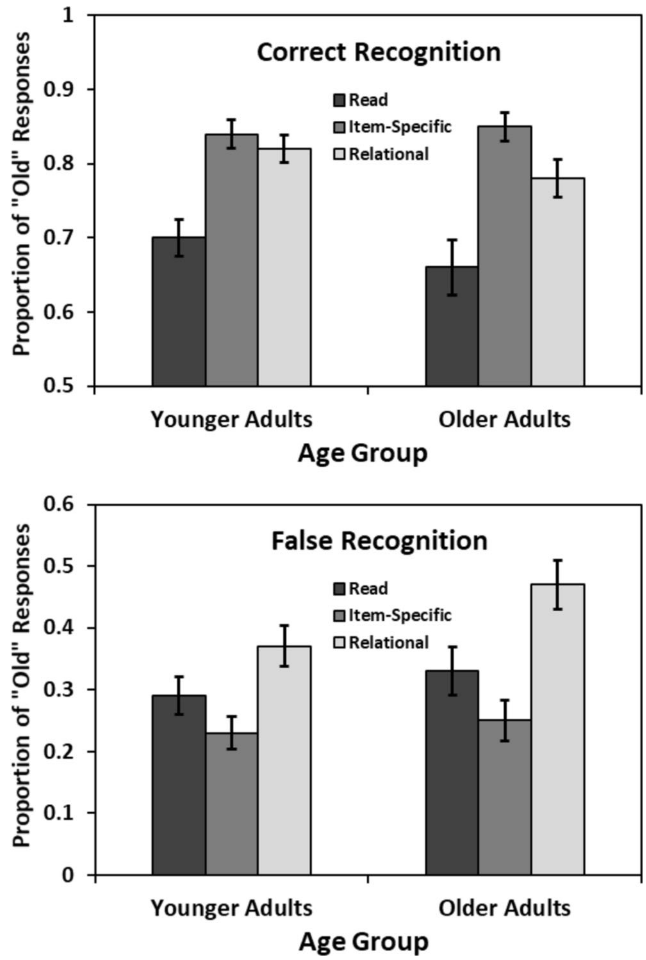

Fig. 2 Proportion of "old" responses to studied list items (top panel) and critical lures (bottom panel) as a function of processing instructions and age group. Error bars are 95% confidence intervals of the means

both the read group (.85 vs. .68), $t(105) = 6.34$, standard error of the mean (SEM) = .02, $d = 1.24$, and the relational group (.85 vs. .80), $t(107) = 2.40$, $SEM = .02$, $d = 0.46$. Further, correct recognition was also greater in the relational group than the read group (.80 vs. .68), $t(108) = 4.30$, $SEM = .02$, $d = 0.83$, effects consistent with deep-processing benefits of item-specific and relational encoding. Both the main effect of age, $F(1, 157) = 1.61$, $MSE = .02$, $p = .21$, and the Encoding × Age interaction, $F < 1$, failed to reach significance, demonstrating that both age groups derived similar correct recognition benefits of item-specific and relational encoding over reading.

We then analyzed signal detection indices to estimate the amount of encoded memory information (d') and amount of memory monitoring at test (lambda) as a function of encoding task and age group. Beginning with encoded memory information (see Fig. 3, top panel), an effect of encoding, $F(2, 157) = 30.84$, $MSE = .46$, $\eta_p^2 = .28$, revealed that memory information was greater in the item-specific group than both the read group (3.02 vs. 1.99), $t(105) =$

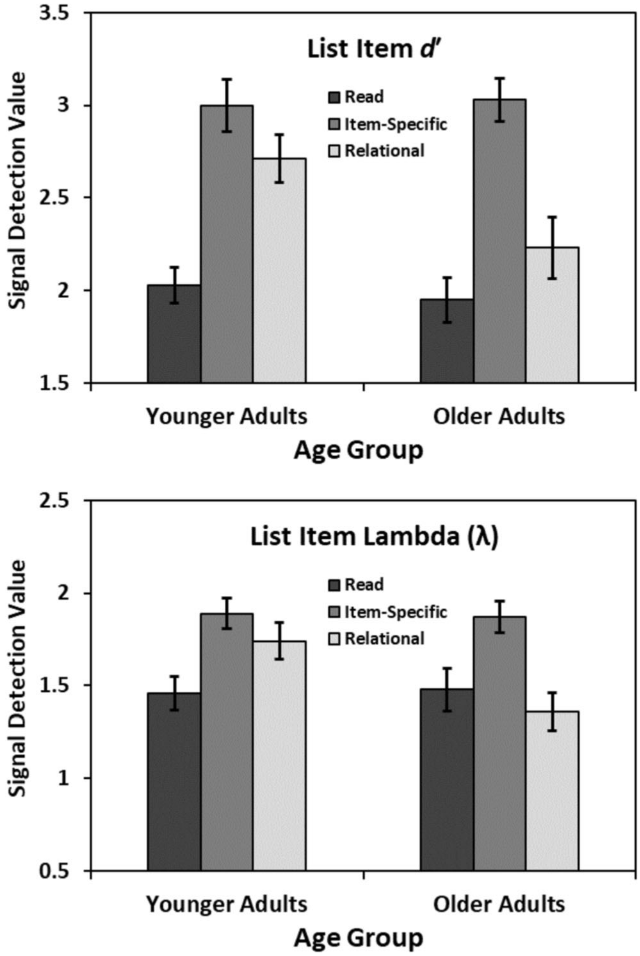

Fig. 3 Mean signal-detection values for list-item d' (top panel) and list-item lambda (λ; bottom panel) as a function of processing instructions and age group. Error bars are 95% confidence intervals of the means

8.65, *SEM* = .08, d = 1.69, and the relational group (3.02 vs. 2.49), $t(107)$ = 3.75, *SEM* = .10, d = 0.73, and was also greater in the relational group than the read group (2.49 vs. 1.99), $t(108)$ = 3.72, *SEM* = .09, d = 0.72. The effect of age was not reliable, $F(1, 157)$ = 2.71, *MSE* = .47, p = .10, nor was the interaction, $F(2, 157)$ = 2.15, *MSE* = .47, p = .12.

Turning to memory monitoring (see Fig. 3, bottom panel), lambda values were also found to differ as a function of encoding task, $F(2, 157)$ = 9.97, *MSE* = .25, η_p^2 = .11. Like d', monitoring was greater following item-specific encoding than both reading (1.88 vs. 1.47), $t(105)$ = 4.43, *SEM* = .06, d = 0.86, and relational encoding (1.88 vs. 1.56), $t(107)$ = 3.28, *SEM* = .06, d = 0.63, but unlike d', monitoring was equivalent between the relational and read groups (1.56 vs. 1.47), $t < 1$. The effect of age was not significant, $F(1, 157)$ = 2.48, *MSE* = .25, p = .11, but a marginal interaction was found, $F(2, 157)$ = 2.62, *MSE* = .25, p = .08, η_p^2 = .03, which reflected a greater decrease in monitoring for older than for younger adults, following relational encoding relative to reading (1.36 vs. 1.74), $t(54)$ = 2.62, *SEM* = .10, d = 0.71, with all other comparisons being nonsignificant, ts < 1.

We then analyzed two diffusion model parameters (drift rate and boundary separation) to determine their convergence with signal-detection estimates of encoded memory information and memory monitoring. Beginning with list-item drift rate (see Fig. 4, top panel), an effect of encoding was found, $F(2, 153)$ = 14.45, *MSE* = .51, η_p^2 = .16, which showed that the drift rate was greater in the item-specific group than the read group (1.58 vs. 0.83), $t(103)$ = 5.44, *SEM* = .10, d = 1.07, and marginally greater than the relational group (1.58 vs. 1.31), $t(105)$ = 1.84, *SEM* = .10, p = .07, d = 0.36. The drift rate was also found to be greater in the relational group than in the read group (1.31 vs. 0.83), $t(104)$ = 3.28, *SEM* = .10, d = 0.64. Unlike our d' estimates, however, a significant effect of age was found, $F(1, 153)$ = 11.27, *MSE* = .51, η_p^2 = .07, which indicated that memory information for list items accumulated more quickly for younger than for older adults (1.43 vs. 1.05). The Encoding × Age interaction was not significant, $F < 1$.

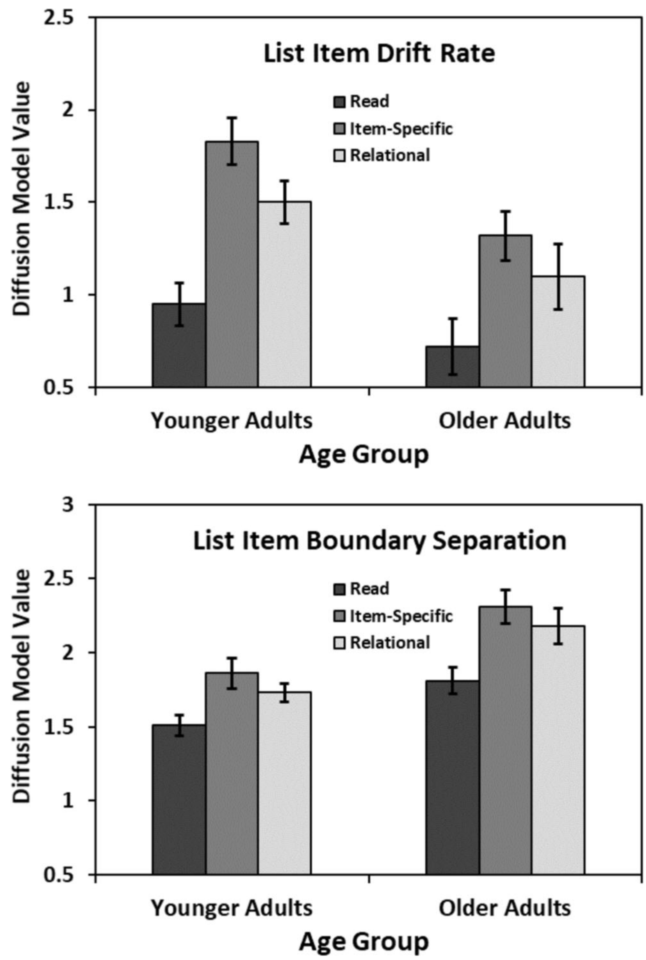

Fig. 4 Mean diffusion model values for list-item drift rate (top panel) and list-item boundary separation (bottom panel) as a function of processing instructions and age group. Error bars are 95% confidence intervals of the means

Response-caution estimates where then analyzed by comparing boundary separation estimates across groups (see Fig. 4, bottom panel). An effect of encoding was found, $F(2, 153) = 10.55$, $MSE = .24$, $\eta_p^2 = .12$, which indicated that response caution was greater following item-specific encoding than it was for reading (2.08 vs. 1.66), $t(103) = 4.14$, $SEM = .07$, $d = 0.82$, and following relational encoding than it was for reading (1.95 vs. 1.66), $t(104) = 3.04$, $SEM = .07$, $d = 0.60$. Response caution was equivalent between the item-specific and relational groups (1.95 vs. 1.66), $t(105) = 1.26$, $SEM = .08$, $p = .21$. An effect of age was also found, $F(1, 153) = 26.57$, $MSE = .24$, $\eta_p^2 = .15$, in which boundary separation was greater in older adults than in younger adults (2.10 vs. 1.70), indicating that older adults showed greater response caution than younger adults did. The Encoding × Age interaction was not significant, $F < 1$.

False recognition

The proportion of falsely recognized critical lures are reported in Fig. 2 (bottom panel), and were analyzed using the same ANOVA. An effect of encoding was found, $F(2, 157) = 13.69$, $MSE = .03$, $\eta_p^2 = .15$. Consistent with item-specific effects in the DRM paradigm, false recognition was lower following item-specific encoding relative to both relational encoding (.24 vs. 42), $t(107) = 5.09$, $SEM = .02$, $d = 0.98$, and reading (.24 vs. .31), $t(104) = 2.07$, $SEM = .02$. Importantly, false recognition was greater in the relational group relative to the read group (.42 vs. .31), $t(108) = 2.92$, $SEM = .02$, $d = 0.56$, a novel pattern (cf. Huff & Bodner, 2013). The effect of age was right at the level of conventional significance, $F(1, 157) = 3.84$, $MSE = .03$, $p = .05$, $\eta_p^2 = .02$, in which false recognition was greater in older adults than in younger adults (.35 vs. .30). The interaction was not reliable, $F < 1$, demonstrating that encoding tasks produced similar effects in both age groups.

Signal-detection analyses were similarly applied to evaluate encoding and monitoring contributions in false recognition. For encoded memory information (see Fig. 5, top panel), an effect of encoding was found, $F(2, 157) = 10.35$, $MSE = .25$, $\eta_p^2 = .12$. Follow-up tests revealed that item-specific encoding significantly reduced the amount of encoded memory information for critical lures relative to relational encoding (1.05 vs. 1.30), $t(107) = 2.59$, $SEM = .07$, $d = 0.51$, a pattern that was similarly found between the read and relational groups (0.87 vs. 1.30), $t(108) = 4.73$, $SEM = .07$, $d = 0.91$. Encoded memory information was marginally lower in the read group than in the item-specific group (0.87 vs. 1.05), $t(105) = 1.88$, $SEM = .07$, $p = .06$, $d = 0.37$. The effect of age and the interaction were not significant, Fs < 1.

For lambda monitoring estimates (Fig. 5, bottom panel), an effect of encoding was found, $F(2, 157) = 7.66$, $MSE = .32$, $\eta_p^2 = .09$, which revealed that monitoring was greater in the item-specific group than both the read group (1.85 vs. 1.44), $t(105) = 3.07$, $SEM = .08$, $d = 0.60$, and the relational group (1.85 vs. 1.55), $t(107) = 2.92$, $SEM = .07$, $d = 0.56$. Monitoring between the relational and read groups was equivalent (1.55 vs. 1.44), $t < 1$. Neither the effect of age, $F(1, 157) = 2.05$, $MSE = .32$, $p = .16$, nor the interaction, $F(2, 157) = 1.19$, $MSE = .32$, $p = .31$, were reliable.

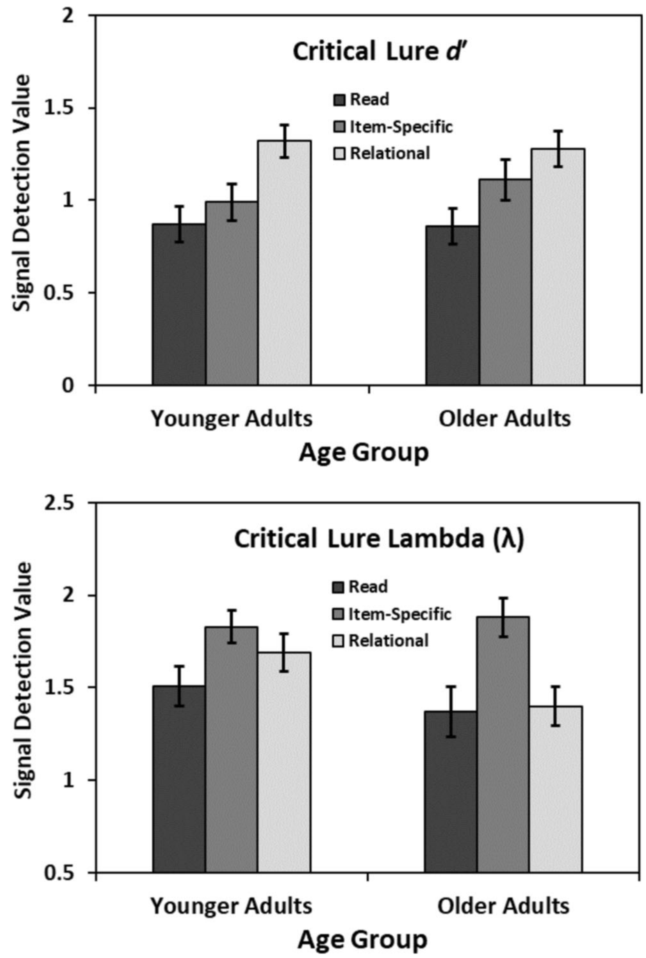

Fig. 5 Mean signal-detection values for critical-lure d' (top panel) and critical-lure lambda (λ; bottom panel) as a function of processing instructions and age group. Error bars are 95% confidence intervals of the means

Diffusion-model parameters were then analyzed for responses to critical lures. When computing drift rates, since most responses to critical lures across participants were "new," the diffusion model provides average drift-rate estimates in the direction of "new" responses (see Fig. 6, top panel), which are negative, indicating that larger values (i.e., being more negative) reflect an accumulation rate that is faster toward making a new response and rejecting the critical lure, and smaller values (i.e., being less negative), reflect an accumulation rate that is slower toward making a new response.

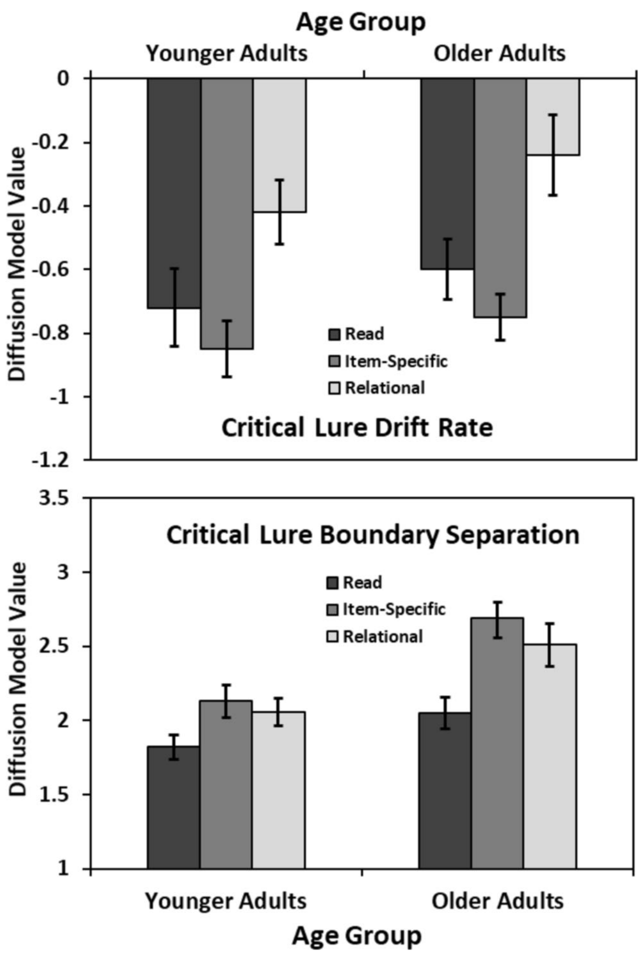

Fig. 6 Mean diffusion model values for critical-lure drift rate (top panel) and critical-lure boundary separation (bottom panel) as a function of processing instructions and age group. Negative drift-rate values reflect that the majority of critical-lure recognition responses were new (vs. old). Error bars are 95% confidence intervals of the means

Our drift-rate analysis again revealed an effect of encoding, $F(2, 153) = 10.56$, $MSE = .29$, $\eta_p^2 = .12$. Follow-up tests revealed that the accumulation of memory evidence to reject critical lures as new was faster in the item-specific than in the relational group (−0.80 vs. −0.33), $t(105) = 4.60$, $SEM = .07$, $d = 0.90$, and similarly faster in the read group relative to the relational group (−0.66 vs. −0.33), $t(104) = 2.84$, $SEM = .08$, $d = 0.56$. There was no difference in drift rates between the item-specific and read groups (−0.80 vs. −0.66), $t(103) = 1.42$, $SEM = .07$, $p = .16$. The effect of age, $F(1, 153) = 2.32$, $MSE = .29$, $p = .13$, and the interaction, $F < 1$, were not reliable.

Response caution for critical lures as estimated through boundary separation was similarly analyzed (see Fig. 6, bottom panel). An effect of encoding was found, $F(2, 153) = 9.26$, $MSE = .34$, $\eta_p^2 = .11$, which revealed that response caution was greater in the item-specific group relative to the read group (2.41 vs. 1.94), $t(103) = 4.03$, $SEM = .08$, $d = 0.79$, and also greater in the relational than in the read group (2.28 vs. 1.94), $t(104) = 2.99$, $SEM = .07$, $d = 0.59$. Response caution was equivalent between the item-specific and relational groups (2.40 vs. 2.28), $t < 1$. As was found for list items, an effect of age was also found, $F(1, 153) = 19.63$, $MSE = .34$, $\eta_p^2 = .11$, in which caution was greater for older than for younger adults. The interaction was not significant, $F(2, 153) = 1.16$, $MSE = .34$, $p = .32$.

General discussion

The purpose of the present experiment was to compare item-specific and relational processing effects to a read-only control on correct and false recognition and to determine whether these effects yield age-related differences. Compared with relational processing and reading, item-specific processing produced a mirror effect pattern: increased correct and decreased false recognition. In contrast, relational processing produced an increase in correct recognition relative to reading, but also produced the expected increase in false recognition. Thus, despite relational processing producing deep semantic encoding which benefitted correct recognition, it compromised memory accuracy. Importantly, equivalent patterns were found in younger and older adults, demonstrating that both age groups show the benefits and costs associated with item-specific and relational processing.

We then estimated the contributions of encoding and monitoring processes that contributed to correct and false recognition. A signal-detection analysis provided d' and lambda, which were used to estimate the quantity of encoded memory information and monitoring, respectively (based on Gunter et al., 2007; Huff & Bodner, 2013). The diffusion model was then used to derive estimates of drift rate and boundary separation to provide separate estimates of encoding and monitoring processes to provide complimentary evidence of these processes.

The application of signal detection and the diffusion model revealed relatively consistent bases for the recognition patterns. For correct recognition, item-specific and relational processing both produced an increase in encoded memory information (through d' and drift rate) and an increase in monitoring processes (through lambda and boundary separation), and these patterns were similar across younger and older adults. For false recognition, relational processing produced an increase in encoded memory information as evidenced by an increase in d' and a less negative drift rate, indicating that relational processing decreased the rate in which critical lures were rejected. The reduced drift rate following relational processing suggests that participants had to overcome elevated levels of encoded memory information to successfully reject critical lures as new. For monitoring, item-specific and relational processing generally increased monitoring relative to reading as indicated by increased lambda and boundary separation estimates. These findings demonstrate that item-specific processing reduces false recognition relative to relational processing by disrupting encoding of implicit activation of critical lures (i.e., impoverished relational encoding; Hege & Dodson, 2004) and by increasing monitoring consistent with the distinctiveness heuristic (Schacter et al., 1999). Monitoring was similarly elevated following relational processing but failed to reduce false recognition, demonstrating that monitoring, particularly following relational processing, may be ineffective at reducing false recognition. When compared with the read group, however, item-specific processing only produced an increase in monitoring, but not encoding, which may explain why item-specific processing yielded only a modest reduction in false recognition.

Although our signal-detection and diffusion-model analyses showed close correspondence, we do note one discrepancy between the estimates. Specifically, for correct recognition drift rate and boundary separation, we found that both diffusion model indices showed an effect of age, whereas the signal-detection indices did not. This age effect occurred despite all encoding task effects producing similar effects across encoding groups. This finding may suggest that the diffusion model is more sensitive to age-related differences due to the use of response latencies versus accuracy to compute the parameters. Consistent with other speeded tasks, such as Stroop's color naming, we note that age-related differences in accuracy are sometimes small between younger and older adults, but differences in response latencies can be large, even when controlling for age-related slowing (Spieler, Balota, & Faust, 1996). Thus, age differences found between signal-detection and diffusion-model estimates may reflect the differences in the sensitivity of accuracy and response latencies in revealing age differences. Indeed, the age effects found in our data set (reduced correct recognition drift rate and greater boundary separation in older adults) have similarly been reported in other recognition analyses that did not use associative lures (Ratcliff et al., 2004). Yet, despite these main effects, the patterns found across the different encoding tasks were quite consistent (as evident by null interactions across analyses), providing convergent evidence for the use of both analyses in examining encoding and retrieval processes across processing types.

Our experiment stems from Huff and Bodner (2013), who similarly compared item-specific and relational processing to a read control using the DRM paradigm. We extend this previous research with several novel additions including a comparison between younger and older adults, the use of categorized word lists and critical lures, and the inclusion of diffusion modeling to separate encoding and retrieval contributions. Despite these differences, the data patterns and conclusions were largely similar, providing greater generalization of this early work. There was, however, a notable difference. Specifically, Huff and Bodner did not find an increase in false recognition following relational processing over reading, whereas our younger and older adult relational processing groups did show an increase in false recognition over reading. We suggest that this difference may be due to the detectability of lures in Huff and Bodner's DRM lists and our categorized lists. In DRM lists, all items converged to a single lure, whereas categorized list items were related to, but did not directly converge on, five different lures. Evidence for differences in lure detection have been illustrated in two prior studies conducted by the first author. In these studies, participants studied DRM lists (Coane, Huff, & Hutchison, 2016) or categorized lists (Huff et al., 2016) constructed identically to the present lists and were asked to explicitly guess what the critical lure(s) was after study. Participants were more likely to correctly guess DRM lures (.36) than categorized lures (ranging from .20 to .23 across experiments), suggesting that DRM lures are indeed easier for participants to generate and possibly detect on a later test. Given these detection rates, relational processing may therefore facilitate detection of DRM lures at study, reducing the likelihood that relational processing would inflate false recognition over reading.

Our study further sought to evaluate the effectiveness of item-specific processing in younger and older adults. As mentioned above, previous literature has been mixed regarding the benefits of item-specific processing in older adults. In our study, older adults showed a mirror effect following item-specific processing, which held when compared with either the read or relational processing groups. The item-specific benefit to false recognition is particularly noteworthy because this pattern occurred despite older adult's greater false recognition rates overall compared with younger adults. Older adults reduced false recognition following item-specific processing contrasts with Butler et al. (2010), who reported elevated false recall rates for older adults using item-specific instructions compared with a read control. We argue that this discrepancy may reflect differences in how effectively participants were utilizing item-specific processing. Specifically, Butler et al. instructed participants to state each item-specific characteristic aloud during study, but reported that older adults would often list characteristics that were convergent with the list theme rather than unique, suggesting use of relational processing. In contrast, our item-specific participants completed a practice list aloud with feedback to discourage the use of relational characteristics at study. Further, Butler et al. used strongly associated DRM lists compared with our categorized lists, which may have been more difficult to promote unique characteristics at study and used free recall rather than recognition. Several differences may therefore contribute to the discrepancies, though our experiment suggests that older adults show similar benefits as younger adults and through similar mechanisms, provided item-specific processing is effectively deployed at study.

Our study also highlights the importance of both signal detection and the diffusion model for providing estimates of encoding and retrieval processes that are often of interest to researchers. These analyses, either individually or in tandem, are advantageous over other methods such as inclusion instructions or within-subjects designs because they allow for the possibility that encoding and retrieval processes contribute to item-specific and relational processing effects in recognition, as found in our analyses. Of course, signal detection and diffusion modeling analyses are not beyond reproach. For instance, they do not provide qualitative information about precisely how individuals are encoding items at study or how monitoring is conducted at test, and only provide quantitative estimates. Qualitatively, we assume that encoded memory information refers to memory strength for list items when applied to correct recognition, and either the amount of associative activation (Roediger et al., 2001) or thematic gist (Brainerd & Reyna, 2002) encoded for critical items when applied to encoded false recognition. Similarly, we assume that monitoring indices for correct and false recognition refer to diagnostic monitoring processes (e.g., Gallo, 2004, 2010) that allow participants to effectively filter studied versus nonstudied items. Our quantitative estimates do not assess these fine-grained processes, and therefore we are careful to make specific claims.

Future research should examine how quantitative estimates of encoding and monitoring correspond to qualitative processes and should also test the validity of these estimates in other domains. For instance, determining whether estimates of encoding and retrieval are sensitive to manipulations that directly affect these processes would promote validity of the indices. To provide some examples, disrupting attention at encoding (Craik & McDowd, 1987) or utilizing speeded response deadlines at test (Benjamin, 2001) would be expected to decrease encoding and retrieval estimates, respectively, whereas repeating items at study and providing warning instructions at test would be expected to increase them. Carefully evaluating both signal-detection and diffusion-modeling parameters under these conditions would provide a stronger test regarding the validity of these measures.

Further, it should also be clear that our item-specific and relational tasks are not process pure and only bias the relative amount of item-specific or relational processing relative to reading. There is always a concern that a given encoding task does not recruit the anticipated type of processing (e.g., Butler et al., 2010), and therefore we are cautious regarding our processing claims. The consistency of our findings with other experiments (Huff & Bodner, 2013; McCabe et al., 2004; Thomas & Sommers, 2005) provide increased confidence that younger and older adults in our experiment are engaging in the expected processing type in the absence of a direct measure.

Assuming our processing tasks do indeed recruit the expected processing type, one important pattern worth emphasizing is the benefits of item-specific processing in both age groups. Our data revealed that item-specific processing produced a mirror effect relative to reading and performing relational processing (cf. Butler et al., 2010). This pattern is important because it demonstrates that older adults can improve memory accuracy despite potential age-related declines. Although participants were utilizing item-specific processing on categorized lists, a context of similarity that may have made item-specific processing more successful, we believe that the application of item-specific processing to benefit memory for real-world materials may be effective, particularly in older adults. For instance, applying imagery (e.g., Foley et al., 2006) to encode unique visual features of grocery-list items or facial features of newly introduced acquaintances may facilitate memory for those materials. Further, item-specific strategies may also reduce commission errors for related items or names/faces. Of course, future research will need to explore precisely how item-specific processing may be adapted beyond word lists, but the benefits shown by older adults suggest that item-specific processing may be a useful encoding strategy.

In sum, we suggest that signal-detection and diffusion modeling analyses can be useful tools for determining the contributions of encoding and retrieval processes in study tasks and in younger and older adults. Through these analyses, we found that both older and younger adults improve recognition accuracy through item-specific over relational processing and reading, and this improvement reflects a combination of both encoding and retrieval processes. Further, we found that relational processing can inflate false recognition over reading despite elevated rates of memory monitoring. Thus, monitoring may be ineffective following high rates of relational processing.

Author note/disclosure of interest Funding for this study was provided by the Aubrey Keith Lucas and Ella Ginn Lucas Endowment at The University of Southern Mississippi. We thank Anastasia Peach and Alyssa Di Mauro for assistance with data collection.

Appendix

Table 2 Mean (*SD*) proportion of "old" responses signal-detection indices and diffusion-model indices for read, item-specific, and relational instructions in younger and older adults

Item type/index/group	Read group	Item-specific group	Relational group
Younger adults			
N	28	27	30
List items	.70 (.13)	.84 (.10)	.82 (.10)
List-item controls	.09 (.08)	.04 (.04)	.07 (.08)
List-items *d*′	2.03 (.51)	3.00 (.73)	2.71 (.71)
List-items drift rate	0.95 (.61)	1.83 (.66)	1.50 (.64)
List-items λ	1.46 (.48)	1.89 (.42)	1.74 (.54)
List-items boundary sep.	1.51 (.37)	1.86 (.53)	1.73 (.34)
Critical lures	.29 (.16)	.23 (.14)	.37 (.18)
Critical-lure controls	.09 (.09)	.05 (.06)	.07 (.10)
Critical-lures *d*′	0.87 (.50)	0.99 (.51)	1.32 (.49)
Critical-lures drift rate	−0.72 (.65)	−0.85 (.46)	−0.42 (.55)
Critical-lures λ	1.51 (.57)	1.83 (.47)	1.69 (.57)
Critical lures boundary sep.	1.82 (.43)	2.13 (.56)	2.06 (.51)
Older adults			
N	26	26	26
List items	.66 (.19)	.85 (.10)	.78 (.13)
List-item controls	.10 (.11)	.04 (.05)	.11 (.11)
List-items *d*′	1.95 (.61)	3.03 (.59)	2.23 (.84)
List-items drift rate	0.72 (.77)	1.32 (.67)	1.10 (.90)
List-items λ	1.48 (.59)	1.87 (.44)	1.36 (.53)
List-items boundary sep.	1.81 (.46)	2.31 (.59)	2.18 (.60)
Critical lures	.33 (.20)	.25 (.17)	.47 (.20)
Critical-lure controls	.13 (.15)	.05 (.07)	.11 (.09)
Critical-lures *d*′	0.86 (.50)	1.11 (.55)	1.28 (.49)
Critical-lures drift rate	−0.60 (.48)	−0.75 (.37)	−0.24 (.65)
Critical-lures λ	1.37 (.69)	1.88 (.54)	1.40 (.53)
Critical-lures boundary sep.	2.05 (.53)	2.69 (.69)	2.51 (.74)

References

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval younger and older adults: I. Evidence from divided attention costs. *Psychology and Aging, 13*, 405–423. https://doi.org/10.1037/0882-7974.13.3.405

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

Aschenbrenner, A. J., Balota, D. A., Gordon, B. A., Ratcliff, R., & Morris, J. C. (2016). A diffusion model analysis of episodic recognition in preclinical individuals with a family history for Alzheimer's disease: The adult children. *Neuropsychology, 3*, 225–238. https://doi.org/10.1037/neu0000222

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., III, McDermott, K. B., & Yerys, B. E. (1999). Veridical false memory in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology, 16*, 361–384. https://doi.org/10.1080/026432999380834

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.), Handbook of memory (pp. 395–410). Oxford, UK: Oxford University Press.

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut norms. *Journal of Experimental Psychology, 80*, 1–46. https://doi.org/10.1037/h0027577

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

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164–169. https://doi.org/10.1110/1467-8721.00192

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

Coane, J. H., Huff, M. J., & Hutchison, K. A. (2016). The ironic effect of guessing: Increased false memory for mediated lists in younger and older adults. *Aging, Neuropsychology, and Cognition, 23*, 282–303.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits. In F. I. M. Craik & S. Trehub (Eds.), Aging and cognitive processes (pp. 191–211). New York, NY: Plenum.

Craik, F. I. M, & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 474–479. https://doi.org/10.1037/0278-7393.13.3.474

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

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavioral Research Methods, Instruments and Computers, 28*, 1–11. https://doi.org/10.3758/BF03203630

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. https://doi.org/10.1002/acp.1265

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

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. https://doi.org/10.1037/0278-7393.30.1.120

Gallo, D. A., Bell, D. M., Beier, J. S., & Schacter, D. L. (2006). Two types of recollection-based monitoring in younger and older adults: Recall-to-reject and the distinctiveness heuristic. *Memory, 14*, 720–741. https://doi.org/10.1080/09658210600648506

Gallo, D. A. (2006). Associative illusions of memory: False memory research in DRM and related tasks. New York, NY: Psychology Press.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition, 38*, 833–848. https://doi.org/10.3758/MC.38.7.833

Gallo, D. A., Roediger, H. L., III, & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review, 8*, 579–586. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1037/0278-7393.30.4.787

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

Huff, M. J., Balota, D. A., & Hutchison, K. A. (2016). The costs and benefits of testing and guessing on recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1559–1572. https://doi.org/10.1037/xlm0000269

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. https://doi.org/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. https://doi.org/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. https://doi.org/10.3758/s13423-014-0648-8

Huff, M. J., & Hutchison, K. A. (2011). The effects of mediated word lists on false recall and false recognition. *Memory & Cognition, 39*, 941–953. https://doi.org/10.3758/s13421-011-0077-0

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. https://doi.org/10.1093/acprof:oso/9780195169669.003.0001

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

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

Hunt, R. R., & Smith, R. E. (2014). How distinctive processing enhances hits and reduces false alarms. *Journal of Memory and Language, 75*, 45–57. https://doi.org/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. https://doi.org/10.1016/j.jml.2011.06.003

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

Kausler, D. H. (1994). Learning and memory in normal aging. San Diego, CA: Academic Press.

Kelley, M. R., & Nairne, J. S. (2001). Von Restoff revisited: Isolation, generation, and memory for order. *Journal of Experimental Psychology, 27*, 54–66. https://doi.org/10.1037/0278-7393.27.1.54

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. https://doi.org/10.3758/BF03210820

Macmillan, N. A., & Creelman, C. D. (1991). Detection theory: A user's guide. New York, NY: Cambridge University Press.

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. https://doi.org/10.3758/BF03196739

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

McCabe, D. P., & Smith, A. D. (2006). The distinctiveness heuristic in false recognition and false recall. *Memory, 14*, 570–583. https://doi.org/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. https://doi.org/10.1037/0096-3445.109.1.49

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

Meade, M. L., & Roediger, H. L., III. (2006). The effect of forced recall on illusory recollection in younger and older adults. *The American Journal of Psychology, 119*, 433–462. https://doi.org/10.2307/120445352

Miller, M. B., & Wolford, G. L. (1999). Theoretical commentary: The role of criterion shifts in false memory. *Psychological Review, 106*, 398–405. https://doi.org/10.1037/0033-295X.106.2.398

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

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*, 838–848. https://doi.org/10.3758/BF03211328

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

Pierce, B. H., Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2005). The modality effect in false recognition: Evidence for test-based monitoring. *Memory & Cognition, 33*, 1407–1413. https://doi.org/10.3758/BF03193373

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language, 50*, 408–424. https://doi.org/10.1016/j.jml.2003.11.002

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology, 60*, 127–157. https://doi.org/10.1016/j.cogpsych.2009.09.001

Robin, F. (2010). Imagination and false memories. *Imagination, Cognition and Personality, 30*, 407–424. https://doi.org/10.2190/IC.30.4.e

Roediger, H. L., III, Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In Roediger H. L. 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. https://doi.org/10.1037/10394-006

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. https://doi.org/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. https://doi.org/10.1006/jmla.1998.261

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

Shipley, W. C. (1986). Shipley Institute of Living Scale. Los Angeles, CA: Western Psychological Services.

Smith, R. E. (2006). Adult age differences in episodic memory: Item-specific, relational, and distinctive processing. In R. R. Hunt & J. B. Worthen (Eds.), Distinctiveness and memory (pp. 259–287). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169669.003.0012

Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5*, 710–715. https://doi.org/10.3758/BF03208850

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. https://doi.org/10.1037/pag0000044

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 461–479. https://doi.org/10.1037/0096-1523.22.2.461

Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition, 29*, 424–432. https://doi.org/10.3758/BF03196393

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.

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

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging, 13*, 230–241. https://doi.org/10.1037/0882-7974.13.2.230

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*, 289–335. https://doi.org/10.1016/j.jml.2003.10.003

von Restorff, H. (1933). Uber die wirkung von bereichsbildungen in spurenfeld [About the effect of area formations in the trace field]. *Psychologische Forschung, 18*, 299–342. https://doi.org/10.1007/BF02409636

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavioral Research Methods, 39*, 767–775. https://doi.org/10.3758/BF03192967

Wahlheim, C. N., & Huff, M. J. (2015). Age differences in the focus of retrieval: Evidence from dual-list free recall. *Psychology and Aging, 30*, 768–780. https://doi.org/10.1037/pag0000049

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorder. *Journal of Mathematical Psychology, 54*, 39–52. https://doi.org/10.1016/j.jmp.2010.01.004

Wixted, J. T., & Stretch, V. (2000). The case against a criterion shift account of false memory. *Psychological Review, 107*, 368–376. https://doi.org/10.1037/0033-295.X.107.2.368