Psychology and Aging
2013, Vol. 28, No. 3, 754–767

0882-7974/13/$12.00 DOI: 10.1037/a0032945

Older and Wiser: Older Adults' Episodic Word Memory Benefits From Sentence Study Contexts

Laura E. Matzen
Sandia National Laboratories, Albuquerque, NM

Aaron S. Benjamin
University of Illinois at Urbana-Champaign

A hallmark of adaptive cognition is the ability to modulate learning in response to the demands posed by different types of tests and different types of materials. Here we evaluate how older adults process out-of-context words and sentences differently by examining patterns of memory errors. In two experiments, we explored younger and older adults' sensitivity to lures on a recognition test following study of words in these two types of contexts. Among the studied words were compound words such as "blackmail" and "jailbird" that were related to conjunction lures (e.g., "blackbird") and semantic lures (e.g., "criminal"). Participants engaged in a recognition test that included old items, conjunction lures, semantic lures, and unrelated new items. In both experiments, younger and older adults had the same general pattern of memory errors: more incorrect endorsements of semantic than conjunction lures following sentence study and more incorrect endorsements of conjunction than semantic lures following list study. The similar pattern reveals that older and younger adults responded to the constraints of the two different study contexts in similar ways. However, although younger and older adults showed similar levels of memory performance for the list study context, the sentence study context elicited superior memory performance in the older participants. It appears as though memory tasks that take advantage of greater expertise in older adults—in this case, greater experience with sentence processing—can reveal superior memory performance in the elderly.

Keywords: aging, false memory, context effects, conjunction errors, semantic errors

Research on the effects of aging on memory typically paints a rather bleak picture. Older adults often perform poorly on memory tasks relative to younger adults (e.g., Light, 1996). However, although many cognitive abilities decline with age, verbal knowledge increases (Park et al., 2002), as does experience. Memory tests involving linguistic or general-knowledge stimuli thus pit an age-related decrease in memory fidelity against an age-related increase in domain expertise (cf. Fraundorf, Watson, & Benjamin, 2012). In this article, we examine how sentence and list study contexts differentially influence orthographic and semantic memory errors, with the idea that the types of errors promoted by a specific study context reveal something about the nature of how the words are processed during encoding (e.g., Roediger, Weldon & Challis, 1989; Tulving & Thompson, 1973). In the current studies, older and younger adults exhibited parallel effects of study contexts—thus, indicating roughly equivalent encoding effects in the two groups. A somewhat stunning consequence of this in our task was that older adults actually exhibited superior memory to younger adults for words studied in sentences. It seems that similar encoding processes in combination with the older adults' greater experience with language led to better memory performance in the old than in the young.

Memory Errors and Encoding

Memory errors provide a rich source of data for investigating the structure and organization of human memory. In this study, we investigate the effects of two common types of study contexts (sentences and word lists) on two types of verbal memory errors (conjunction errors and semantic errors). Conjunction errors occur when participants mistakenly endorse test words that are perceptually or phonetically recombined versions of actually studied words (Jones & Jacoby, 2001; Reinitz, Lammers, & Cochran, 1992; Underwood, Kapelak, & Malmi, 1976; Underwood & Zimmerman, 1973). The study stimuli in such experiments are usually compound words such as "blackmail" and "jailbird," and the critical lures at test are rearrangements of those words, such as "blackbird." The semantic overlap between the studied words and the conjunction lure is typically very low. Thus, participants endorse conjunction lures when they remember some information about the surface forms of the studied words but relatively little information about their meaning.

In contrast, semantic errors occur when participants remember the semantics of studied words but relatively little about their surface forms. Studies investigating semantic lures have used

This article was published Online First July 8, 2013.

Laura E. Matzen, Albuquerque, NM, Cognitive Systems Research and Applications, Sandia National Laboratories, Aaron S. Benjamin, Psychology Department, University of Illinois at Urbana-Champaign.

This work was funded in part by Grant R01 AG026263 to Aaron Benjamin from The National Institutes of Health. It was also supported by the Sandia National Laboratories Excellence in Engineering Fellowship provided to Laura Matzen and is a part of her doctoral dissertation.

Correspondence concerning this article should be addressed to Laura E. Matzen, Cognitive Systems Research and Applications, Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185-1011. E-mail: lematze@sandia.gov

sentences or stories (Bransford & Franks, 1971; Brewer, 1977; Johnson, Bransford, & Solomon, 1973), lists of semantic associates (Deese, 1959; Roediger & McDermott, 1995), and synonyms (Matzen & Benjamin, 2009; Matzen, Taylor, & Benjamin, 2011) as studied items. At test, participants often falsely recall or recognize words that are semantically similar to the studied items. These findings indicate that participants seem to extract the basic meaning or gist of the studied information and lose information about the surface forms of the words or sentences. This claim is further supported by evidence indicating a role for gist, often in combination with inferences and world knowledge, in reconstructing meaningful text information (Bock & Brewer, 1974; Brewer, 1977; Johnson et al., 1973; Potter & Lombardi, 1990).

Matzen and Benjamin (2009) investigated the effects of study context on different types of memory errors. Participants studied out-of-context words in a list and words in sentence contexts. They were then tested on old items, conjunction lures, semantic lures, and unrelated new items. The results showed an interaction between study context and the prevalence of different types of memory errors. When participants studied sentences, they made more semantic errors and fewer conjunction errors. When they studied words in a list, the pattern reversed and participants made more conjunction errors and fewer semantic errors. These patterns suggest that participants were encoding relatively different types of information from the words in the different study contexts.

When reading sentences, people typically extract the gist of the sentence and do not retain much information about the surface forms of the sentence, such as the syntactic structure or the specific words that were used (Potter & Lombardi, 1990). This is an effective strategy in everyday life because memory for the core meaning of a sentence or story is typically more important and more useful than memory for the surface forms. Participants in an experiment are likely to approach sentences in the same way, particularly if they are experienced readers. If participants remember the gist of a studied sentence but not the specific words that were used, they become more susceptible to semantic lures at test. Conversely, gist-based encoding makes people less susceptible to conjunction lures, which share the surface forms of studied items but do not share their meaning. If information about the surface forms of the words has been lost, conjunction lures will not be appealing. In addition, if the lures do not match with the gist of any of the studied items, they are easier to reject.

A list of words provides a different context and constrains encoding in different ways. In daily life, people often have different goals when they are trying to remember a list of words than when they are trying to comprehend a set of sentences. For example, if someone is trying to remember a grocery list, extracting the gist of that list, "food," is typically not very useful upon reaching the grocery store. With grocery lists, to-do lists, and other types of lists, the specific items are more meaningful than the overall gist of the list. It is also difficult to extract any gist information from a list of unrelated words. When people encounter out-of-context words in a memory experiment, they are likely to bring their real-world experiences to bear and to encode relatively more information about the surface forms of the words and relatively less information about their gist. As described above, this approach would influence their susceptibility to different types of memory lures at test. In this case, conjunction lures would be more appealing and semantic lures would be less so.

The balance between encoding gist versus verbatim details of verbal information is likely more nuanced than the extremes used for illustration here. Alba and Hasher (1983) showed, for example, that people sometimes encode very detailed information about verbal stimuli, including both semantic and surface form information. The information that they access on a memory test is influenced by many factors, including task demands, response biases, and the affordances of the study and test materials. In this article and in other recent work (Matzen & Benjamin, 2009), we manipulate the manner of encoding with a change in study context, though we recognize that many factors influence this balance.

Memory Errors and Aging

Matzen and Benjamin (2009) demonstrated these interactions between study context and error rates for younger adults. However, it is unclear whether or not the same pattern would hold for older adults. Previous research on memory performance has indicated that older adults rely more on gist-based processing than do younger adults (Balota et al., 1999; Kensinger & Schacter, 1999; Norman & Schacter, 1997; Schacter, Koutstaal, Johnson, Gross, & Angell, 1997; Tun, Wingfield, Rosen, & Blanchard, 1998; Watson, Balota, & Sergent-Marshall, 2001). This could make older adults generally more susceptible to semantic errors and less susceptible to conjunction errors than younger adults. The existing research on memory errors in older adults bears out this prediction. Studies using the Deese-Roediger-McDermott (DRM) paradigm with older adults have found that they are more susceptible to semantic memory errors than younger adults (cf. Balota et al., 1999; Benjamin, 2001; Norman & Schacter, 1997; Tun et al., 1998; Watson et al., 2001). However, in experiments using form-based memory lures, older adults' error rates in response to feature and conjunction lures are often quite similar to those of younger adults, despite poorer overall memory performance for old words (Jones & Jacoby, 2005; Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996; Rubin, Van Petten, Glisky, & Newberg, 1999).

There is also considerable evidence that older adults have poor source memory relative to younger adults (cf. Benjamin, 2010; Benjamin, Diaz, Matzen, & Johnson, 2012; Ferguson, Hashtroudi, & Johnson, 1992; Henkel, Johnson, & De Leonardis, 1998; Jones & Jacoby, 2005; McIntyre & Craik, 1987; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Spencer & Raz, 1995) which may make them less able to use source-based information to avoid memory errors. Previous studies using both semantic and conjunction lures have found that older adults are ineffective relative to younger adults at recalling source information and using it to avoid memory errors (Jones & Jacoby, 2005; Pierce, Sullivan, Schacter, & Budson, 2005). Pierce, Sullivan, Schacter, and Budson (2005) investigated older adults' ability to use source memory to avoid gist-based memory errors. They found that when older adults studied words in sentence contexts, they were able to use contextual information to avoid false alarms to a limited extent. However, the older adults were much less successful at using the sentence context information than were younger adults. Similarly, Jones and Jacoby (2005) found that although young adults were able to use information about study modality to reject conjunction lures, older adults were unable to do so.

Other manipulations that reduce rates of false memory in younger adults, such as repeated study trials and warnings about the nature of the lures, typically do not reduce false memories in older adults (Benjamin, 2001; Jacoby, 1999; Kensinger & Schacter, 1999; Watson, McDermott, & Balota, 2004). All of these findings are consistent with a larger body of literature suggesting that older adults do not modulate their encoding in response to study circumstances in the same way that younger adults do. This lack of sensitivity may underlie the finding that training interventions with older adults to improve memory are often ineffective (cf. Anschutz, Camp, Markley, & Kramer, 1987; Bissig & Lustig, 2007; Verhaeghen, Marcoen, & Goossens, 1992; see also Hertzog & Dunlosky, 2005).

Taken as a whole, this collection of results suggests that older adults might not reveal the same pattern of results that was seen for the younger adults in Matzen and Benjamin (2009). If older adults rely more on gist-based processing in general and are unable to modulate their encoding in response to the constraints of different study contexts, then they would be likely to have relatively high rates of semantic errors and low rates of conjunction errors, and this pattern should be minimally affected by study context. Additionally, because older adults generally have poor source memory, they would be unable to use information from the sentence contexts to help them to avoid either type of memory error, thus leading to higher rates of false memory in general.

However, other research suggests that older adults demonstrate flexibility in encoding that can benefit memory performance under the right circumstances. For example, studies investigating older adults' use of memory strategies have shown that older adults can make effective use of memory strategies when those strategies are naturalistic and applicable to real-life tasks outside of the laboratory (Castel, 2008; Dunlosky & Hertzog, 1998; Fraundorf et al., 2012; Rendell & Craik, 2000; Tentori, Osherson, Hasher, & May, 2001) or when the information to be remembered is perceived as being valuable (Castel, Benjamin, Craik, & Watkins, 2002; Halamish, McGillivray, & Castel, 2011; Rahhal, May, & Hasher, 2002). The findings of Matzen and Benjamin (2009) suggest that manipulating the context in which studied words appear changes readers' encoding biases in a very naturalistic way that is likely shaped by their day-to-day experience with lists and texts. Our goal in the present study was to determine whether or not older adults can shift between these two approaches to encoding in the same way that younger adults do. We hypothesized that older adults would exhibit the same pattern of memory errors as younger adults because the effects of these study contexts on encoding are likely due to readers' experiences with similar materials under natural circumstances in their daily lives.

Experiment 1

Method

Participants. Twenty-five older adults (11 female) from the Urbana-Champaign community were paid for their participation in the experiment. The mean age of the participants was 72 (range 60–81). The participants were drawn from a large pool of volunteers over the age of 60 with normal scores on the Mini-Mental State Examination (MMSE; mean score = 29).

Design. The critical variable was whether items were studied within the context of sentences or as isolated words. The item types at test were old (previously seen) words, new (unrelated) words, conjunction and semantic lures whose parent words appeared in sentence contexts at study, and conjunction and semantic lures whose parent words appeared as single out-of-context words at study. All of the critical variables were manipulated within-subjects. The dependent variable was confidence in the recognition judgment, used to generate individual isosensitivity functions within each condition as well as hit and false alarm rates.

Materials. The materials and procedure of Experiment 1 were identical to those used with younger adults in Experiment 3 of Matzen and Benjamin (2009). The experiment used a total of 160 parent compound words (such as "tailspin" and "floodgate") that could be recombined to form conjunction lures ("tailgate"). Of the 160 parent compound words, a subset of 64 parent words that had a synonym or close semantic associate were assigned to the semantic lure set. These parent words and their semantic associates were interchangeable in a sentence context. For example, the semantic associate for "tailspin" was "nosedive," and the two words were both appropriate in the sentence context "The fighter plane went into a tailspin/nosedive after it was hit by enemy fire." Approximately half of the semantic associates were compound words.

The stimuli were divided into eight counterbalanced study lists containing 160 items each. Half of the items on each study list were out-of-context compound words and half were sentences containing a compound word. Each list contained 96 items that were rotated through four conjunction lure conditions and 64 items that were rotated through four semantic lure conditions. The 96 items in the conjunction lure subset of each study list consisted of 64 items (32 sentences and 32 single words) that were recombined to create conjunction lures at test and 32 items (16 sentences and 16 single words) that were used as old items at test. The items were counterbalanced so that each parent word appeared in all four conjunction lure conditions: presented in a sentence and tested as an old item, presented as a single word and tested as an old item, presented in a sentence and combined with a parent word from another sentence to form a lure at test, or presented as a single word and combined with another single parent word to form a lure at test. The semantic lure subset of each study list had 32 items (16 sentences and 16 single words) that contained one member of a pair of close semantic associates. The other member of this pair was presented at test as a to-be-rejected semantic lure. The remaining 32 items in this subset (16 sentences and 16 words) contained one member of a pair of semantic associates that was presented in the same form at test, serving as a to-be-endorsed old item. The semantic items were counterbalanced so that each word (compound parent word or its semantic associate) appeared in all four conditions: presented in a sentence and tested as an old item, presented as a single word and tested as an old item, presented in a sentence with the other member of the pair tested as a lure, or presented as a single word with the other member of the pair tested as a lure.

The 160 study items for each list were placed in a pseudorandom order with the appropriate versions of each item placed in each slot to create eight unique study lists. Each study list had an associated test list that contained 192 items. Of the test items, 32 were conjunction lures, 32 were semantic lures, 64 were old items, and 64 were new, unrelated items. All of the conjunction lures and approximately half of the semantic lures were compound words, so a similar pattern was created for the old and new items in which approximately three fourths of the old and new items were compound words and one fourth were not. The same 64 new items were used for all eight lists. The new items were matched as closely as possible to the old items and lures in terms of length and frequency. The average length of the words on the test list was 8.26 letters for old items, 8.04 letters for lures, and 8.27 letters for new items. The average frequency of the test items was 12.53 for the old items, 10.55 for the lures, and 6.30 for the new items (based on the Kucera & Francis, 1967 norms included in Balota et al., 2002; a frequency value of zero was assumed for items not appearing in the database). The 192 test items for each list were placed in a pseudorandom order so that no more than three items of the same type appeared in a row. The same order was used for all eight test lists with the appropriate test items substituted into each slot.

Experiment 1 used a single study phase followed by a single test list. For each test list, care was taken to ensure that the two morphemes in each conjunction lure appeared the same number of times (twice for one item, once for all other items) in the preceding study list. Additionally, none of the morphemes in any of the semantic lures or new items appeared anywhere in the preceding study list.

Procedure

Participants were instructed that they would be studying a list of intermixed words and sentences for a subsequent memory test. During the study phase, one item at a time (a single word or a sentence) was presented on the computer monitor in black 36-point Times New Roman font on a white background. Single words were presented for two seconds and sentences were presented for eight seconds with a 250 ms interstimulus interval. The words and sentences were quasi-randomly intermixed with no more than four single words or four sentences appearing in a row. During the test phase, participants saw one word at a time and were asked to respond by pressing the keys 1–4 on the computer keyboard. A response of "1" indicated that the participants were sure that the word had *not* appeared on the study list, a response of "2" indicated that they thought the word was new but were not sure, a response of "3" indicated that they thought they had studied the word but were not sure, and a response of "4" indicated that they were sure that they *had* studied the word. Each word stayed on the screen along with a guide indicating what each response choice meant until the participant selected his or her response.

Analysis

To analyze the false-alarm rates (FAR), the number of "yes" responses for each lure condition for each participant was entered into a within-subjects ANOVA with lure type (conjunction vs. semantic) and study context (sentence vs. word) as independent variables. The critical test was whether the FAR was higher for conjunction lures in the list-study condition and higher for semantic lures in the sentence-study condition. That interaction would replicate the basic result seen in younger adults.

This same interaction was sought using the detection-theoretic measure d_a, in order to avoid the measurement concerns that accompany evaluation of proportions on a bounded scale. This measure is based on basic assumptions of the Theory of Signal Detection (Green & Swets, 1966; Macmillan & Creelman, 2005; Peterson, Birdsall, & Fox, 1954; Swets, 1986), as applied to recognition memory (Egan, 1958). Because d_a has metric properties that are suited to comparing memory errors across conditions with different overall levels of performance (see also Banks, 2000; Benjamin, 2005; Benjamin, Diaz, & Wee, 2009; Matzen & Benjamin, 2009), it can be used to make direct comparisons across conditions by subtracting the old-lure d_a value from the old-new d_a value for each participant to generate $\Delta\ d_a$ values (Matzen & Benjamin, 2009). The resulting $\Delta\ d_a$ values indicate how likely the participants are to correctly identify a lure as being a new item. A small $\Delta\ d_a$ indicates that participants responded to lures and to unrelated new items similarly, indicating success at identifying the lures as new words.

Results

Table 1 provides the mean proportions of each confidence rating for each item type. The average hit rates were similar for the old items that were originally studied in sentences ($M = 76\%$) and for old items that were originally studied as single words, 70%, $t(24) = 1.66$, $p = .11$. For the conjunction lures, the average percentage of false alarms was significantly lower in the sentence condition (33%) than in the word condition, 40%, $t(24) = 2.15$, $p < .05$. The opposite was true for the semantic lures, with average false alarm rates that were significantly higher for the sentence condition (40%) than for the word condition, 32%, $t(24) = 3.18$, $p < .01$. A 2 × 2 ANOVA (study context × lure type) revealed a significant interaction between study context (word or sentence) and lure type (conjunction lure or semantic lure; $F(1, 24) = 12.54$, $p < .01$). This pattern replicates the central finding of Matzen and

Table 1
Means and Standard Deviations for Proportions of Each Confidence Rating for Each Item Type in Experiment 1

		Confidence ratings			
Study context	Test item type	1 (Sure new)	2 (Unsure new)	3 (Unsure old)	4 (Sure old)
Sentence condition	Conjunction lures	0.38 (0.22)	0.29 (0.20)	0.15 (0.14)	0.18 (0.17)
	Semantic lures	0.40 (0.25)	0.20 (0.18)	0.13 (0.15)	0.27 (0.18)
	Old words	0.13 (0.10)	0.12 (0.09)	0.11 (0.08)	0.65 (0.16)
Word condition	Conjunction lures	0.35 (0.24)	0.26 (0.20)	0.18 (0.13)	0.22 (0.20)
	Semantic lures	0.46 (0.28)	0.22 (0.19)	0.17 (0.17)	0.16 (0.15)
	Old words	0.13 (0.12)	0.17 (0.12)	0.14 (0.11)	0.56 (0.17)
None	New items	0.50 (0.27)	0.24 (0.18)	0.14 (0.13)	0.12 (0.12)

Benjamin (2009) for younger adults. Table 2 provides the mean proportions of each confidence rating for each item type for the younger adults in Experiment 3 of Matzen and Benjamin (2009). A comparison of the false alarm rates for the older and younger adults is shown in Figure 1.

The same pattern emerged again for the older adults' $\Delta\ d_a$ values, as shown in Figure 2. There was a marginally significant difference in overall discrimination between study context conditions, as shown by the differences in d_a for old-new recognition ($d_a = 1.28$ for word study and $d_a = 1.48$ for sentence study, $t(24) = 1.95$, $p = .06$). Relative to old–new discrimination, there was a bigger decrease in discrimination performance for the conjunction lures whose parents were studied as single words ($d_a = 0.84$, $\Delta\ d_a = 0.44$) than there was for the conjunction lures whose parent items were studied in sentences ($d_a = 1.15$, $\Delta\ d_a = 0.33$). The opposite pattern was obtained for the semantic lures, with a bigger decrease in discrimination for lures whose parent items were presented in sentences ($d_a = 0.97$, $\Delta\ d_a = 0.52$) than there was for lures whose parent items were presented as single words ($d_a = 1.06$, $\Delta\ d_a = 0.22$). The $\Delta\ d_a$ values were significantly different across the two study contexts for the semantic lures, $t(24) = 2.54$, $p < .02$ but not for the conjunction lures, $t(24) = 1.44$, $p = .16$. Similarly, within each study context, there was a significant difference between $\Delta\ d_a$ values for the two types of lures for the list study context, $t(24) = 2.87$, $p < .01$ but not for the sentence study context, $t(24) = 1.60$, $p = .12$ Critically, the interaction between study context (word or sentence) and lure type (semantic lure or conjunction lure) was significant for the $\Delta\ d_a$ values, $F(1, 24) = 10.73$, $p < .01$. This result replicates the finding of Matzen and Benjamin (2009) and indicates that older adults adjust what is encoded from sentences and lists in much the same way that younger adults do, leading to very similar patterns of memory errors for the two groups.

To further explore the differences and similarities between the younger and older adults, we directly compared the older adults' performance with that of the younger adults in Experiment 3 of Matzen and Benjamin (2009). Although the older adults' overall pattern of memory errors qualitatively mirrors that of the younger adults, there are also interesting quantitative differences that emerge from comparisons between the two groups, particularly for the high-confidence responses. The older adults had more high-confidence "yes" responses than the younger adults in every condition. Their high-confidence false alarm rates were significantly higher for unrelated new items, $t(47) = 3.05$, $p < .01$ and for every lure type (all p's < 0.05). By itself, this might imply that the older adults had poorer overall memory performance relative to the younger adults. However, the older adults also had significantly more high-confidence hits than the younger adults for the old words in both the sentence, $t(47) = 4.84$, $p < .01$ and the word condition, $t(47) = 2.50$, $p = .02$. In addition, the older and younger adults did not differ significantly in their number of high-confidence misses for either the sentence condition, $t(47) = 1.43$ or for the word condition, $t(47) = 0.05$. The older and younger adults also had equivalent rates of high-confidence correct rejections for all lure types (all t's < 1.5).

It is difficult to compare the older and younger adults' performance using hit and false alarm rates because of their different patterns of high- and low-confidence hits, misses, and false alarms. However, the d_a values, which incorporate the participants' confidence ratings into a measure of their ability to discriminate between different types of items, provide a much clearer picture. The d_a values (shown for both groups in Figure 3) indicate that the older adults were typically no worse than the younger adults at discriminating old items from new items and lures in the word condition. In the sentence condition, the older adults were actually *better* than the younger adults at making these same discriminations.

The older adults' superior discrimination performance for the sentence study contexts was confirmed in a 2 × 3 × 2 ANOVA (study context × lure type × age), in which there was a significant main effect of lure type, $F(2, 47) = 28.21$, a significant interaction between lure type and study context, $F(2, 47) = 10.69$, and a significant interaction between study context and age, $F(2, 47) = 5.75$. There was not a three-way interaction.

Pairwise comparisons for the simple effects revealed that the older and younger adults' performance was not significantly different for words that were studied out of context (all t's $<$ 1.0). However, for words that were studied in sentences, the older adults had significantly better old–new, $t(47) = 4.24$, $p <$ .01 and old-conjunction lure, $t(47) = 3.54$, $p < .01$ discrimination performance than the younger adults. The older adults' old-semantic lure discrimination performance was numerically better than that of the younger adults, but the difference between the two groups was not significant, $t(47) = 1.49$, $p =$ .15.

Table 2
Means and Standard Deviations for Proportions of Each Confidence Rating for Each Item Type in Experiment 3 of Matzen & Benjamin (2009)

		Confidence ratings			
Study context	Test item type	1 (Sure new)	2 (Unsure new)	3 (Unsure old)	4 (Sure old)
Sentence condition	Conjunction lures	0.35 (0.28)	0.40 (0.22)	0.16 (0.13)	0.09 (0.10)
	Semantic lures	0.35 (0.29)	0.40 (0.27)	0.14 (0.11)	0.11 (0.13)
	Old words	0.21 (0.21)	0.24 (0.16)	0.16 (0.12)	0.39 (0.19)
Word condition	Conjunction lures	0.32 (0.28)	0.39 (0.24)	0.18 (0.16)	0.11 (0.14)
	Semantic lures	0.38 (0.30)	0.40 (0.27)	0.17 (0.12)	0.05 (0.06)
	Old words	0.14 (0.18)	0.25 (0.15)	0.18 (0.13)	0.43 (0.20)
None	New items	0.41 (0.31)	0.42 (0.26)	0.13 (0.12)	0.04 (0.06)

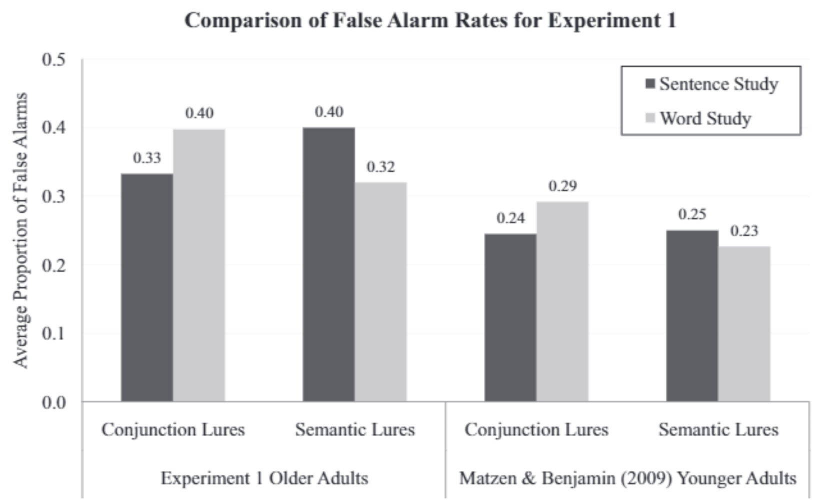

Figure 1. False alarm rates for conjunction lures and semantic lures with parent words that were studied as single words or in sentence contexts. The false alarm rates for the older adults in Experiment 1 are compared with the false alarm rates for younger adults in Experiment 3 of Matzen and Benjamin (2009).

Discussion

Both the false alarm rates and the pattern of d_a values indicate that, just like younger adults, older adults encoded the words differently in response to the two different study contexts. They encoded relatively more gist information for the sentence contexts and relatively more surface information for the decontextualized words. This in turn led the older adults to be more susceptible to semantic lures if the corresponding parent items had been studied in sentences and more susceptible to conjunction lures if the corresponding parent items had been studied out of context. Overall, the results of the experiment support the idea that older adults adapt the way in which they encode information in response to different study contexts in the same qualitative manner as younger adults.

The fact that the d_a values for each item type in the sentence condition were higher for older adults than they were for younger adults is perhaps the most striking finding in this experiment. This indicates that the older adults were better than the young at discriminating new items and lures from old items following sentence study. The older adults' superior discrimination performance in the sentence condition cannot be explained simply by a greater reliance on gist-based processing. If that were the case, we would expect that their discrimination of the semantic lures from old items in the sentence condition would be worse than it was for

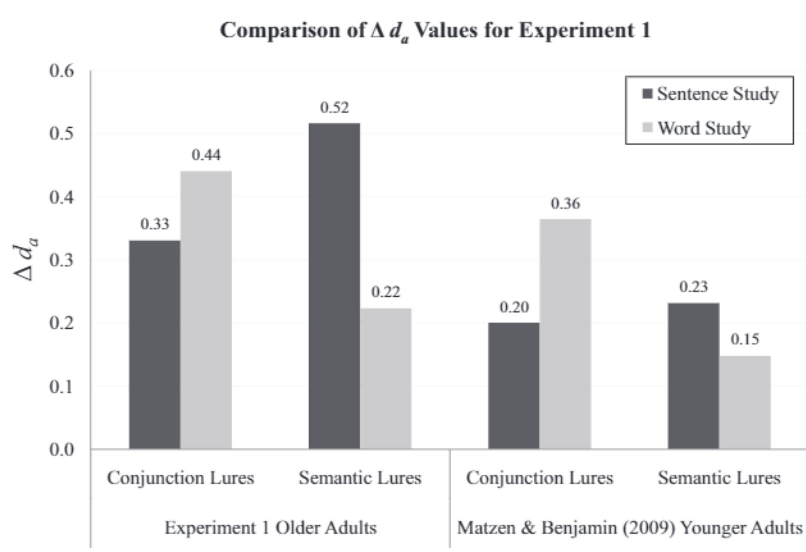

Figure 2. Δ d_a values for Experiment 1 and for Experiment 3 of Matzen and Benjamin (2009). Δ d_a is the difference between the d_a value for old-new discrimination and the d_a value for old-lure discrimination for each condition.

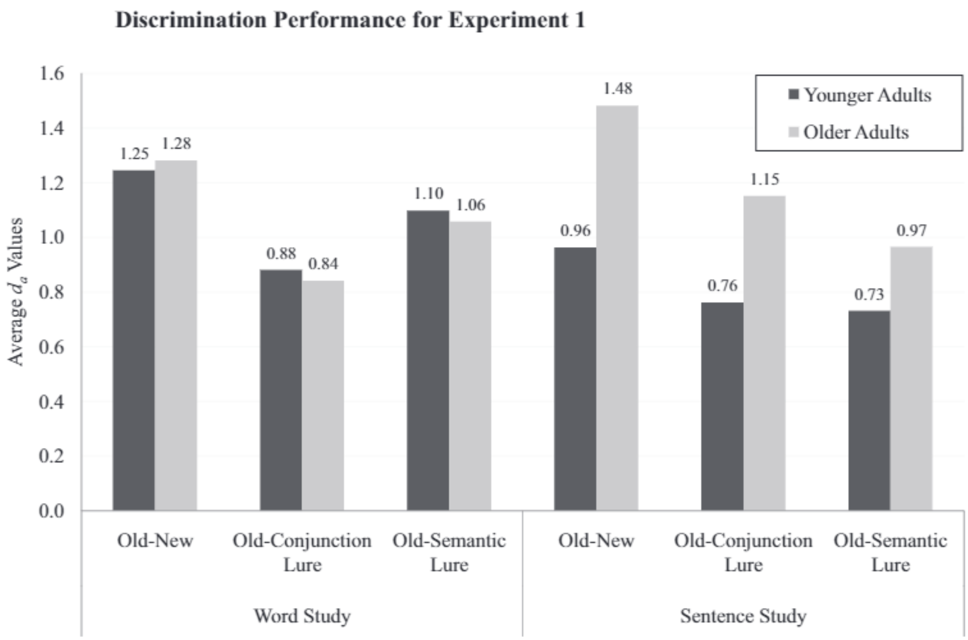

Figure 3. Comparison of discrimination performance (as measured by d_a) for the older adults in Experiment 1 and the younger adults from Experiment 3 of Matzen and Benjamin (2009).

younger adults, yet the opposite is true. A more plausible explanation is that the older adults are retaining more information about the sentence contexts overall, producing both higher hit rates for the old items and enabling recollection-based rejection of some of the lures. Although the older adults had more high-confidence false alarms for the both types of lures in the sentence condition, they also had more high-confidence rejections of those items than the younger adults. If the older adults' higher rates of high-confidence rejections are due to recollection rejection, this study would provide a rare example of older adults using source information more effectively than young adults.

In Experiment 2, we attempted to replicate this surprising finding with new groups of older and younger adults. Experiment 1 used two different subsets of items to produce the conjunction lures and semantic lures, so it is possible that there were unforeseen differences between the item sets. To eliminate this possibility, we devised a fully counterbalanced item set for Experiment 2. We hypothesized that the older and younger adults' susceptibility to the conjunction and semantic lures would vary with study context, replicating the results of Experiment 1 and of Matzen and Benjamin (2009). In addition, we hypothesized that the older adults would show better discrimination performance than the younger adults for the sentence study condition but not for the word study condition.

Experiment 2

Method

Participants. Thirty-two older adults (19 female) participated in the experiment and were paid for their participation. The mean age of the older adult participants was 69 (range 61–82). The participants were drawn from the same pool of older adults described in Experiment 1. The older adults' mean score on the MMSE was 29 (range 26–30). Thirty-two young adults (23 female) participated in the experiment for credit in an introductory psychology course or for payment. The mean age of the young adult participants was 22 (range 18–28). Both groups of participants completed the Shipley Vocabulary Test (Shipley, 1946). The mean score for the older adults was 35 (range 24–39) and the mean score for the younger adults was 32 (range 25–39).

Design. The design of Experiment 2 was identical to Experiment 1. The difference between the two experiments is that Experiment 2 was fully counterbalanced, with all of the parent items appearing in all conditions. In contrast, Experiment 1 used three sets of items: one set of parent items that rotated through the conjunction lure conditions, a second set of parent items that was rotated through the semantic lure conditions, and a matched set of new items. In Experiment 2, the same items appeared in every condition across the experimental lists.

Materials and procedure. Experiment 2 used a set of 256 compound parent words. All of the parent words were in pairs whose parts could be recombined to form conjunction lures and half of the parent words were also paired with a semantic lure. As in Experiment 1, the semantic lures were close associates of the parent word and were interchangeable with that word in a sentence context. The parent words were divided into 16 fully counterbalanced experimental lists. All items were rotated through the 16 counterbalancing conditions illustrated in Table 3. Each item appeared in a different condition in each experimental list. Each experimental list consisted of 192 study items that were presented in a single study phase, followed by 192 test items that were presented in a single test phase. Half of the study items were sentences and half were out-of-context words. The test phase contained 16 conjunction lures and 16 semantic lures whose parent

Table 3
Examples of Items and Counterbalancing Conditions for Experiment 2

Counter-balancing condition	Study condition	Study items	Test condition	Test items
1	Sentence study conjunction order 1	The lumberjack was injured when he couldn't escape the path of a tree. The town's historical society helped to clean up the old graveyard.	Conjunction lure	lumberyard
2	Sentence study conjunction order 2	The town's historical society helped to clean up the old graveyard. The lumberjack was injured when he couldn't escape the path of a tree.	Conjunction lure	lumberyard
3	Word study conjunction order 1	lumberjack graveyard	Conjunction lure	lumberyard
4	Word study conjunction order 2	graveyard lumberjack	Conjunction lure	lumberyard
5	Sentence study semantic order 1	The lumberjack was injured when he couldn't escape the path of a tree. The town's historical society helped to clean up the old graveyard.	Semantic lure	logger cemetery
6	Sentence study semantic order 2	The logger was injured when he couldn't escape the path of a tree. The town's historical society helped to clean up the old cemetery.	Semantic lure	lumberjack graveyard
7	Word study semantic order 1	lumberjack graveyard	Semantic lure	logger cemetery
8	Word study semantic order 2	logger cemetery	Semantic lure	lumberjack graveyard
9	Sentence study	The lumberjack was injured when he couldn't escape the path of a tree. The town's historical society helped to clean up the old graveyard.	Old	lumberjack graveyard
10	Sentence study	The logger was injured when he couldn't escape the path of a tree. The town's historical society helped to clean up the old cemetery.	Old	logger cemetery
11	Word study	lumberjack graveyard	Old	lumberjack graveyard
12	Word study	logger cemetery	Old	logger cemetery
13	N/A	N/A	New	lumberjack graveyard
14	N/A	N/A	New	logger
15	N/A	N/A	New	cemetery
16	N/A	N/A	New	lumberyard

words were studied in sentences, 16 conjunction lures and 16 semantic lures whose parents were studied as out-of-context words, 32 old words that had been studied in sentences, 32 old words that had been studied as out-of-context words, and 64 new words that were not studied in that experimental list. Of the 64 words that were tested as new items in a particular experimental list, 32 were parent items from other lists, 16 were conjunction lures from other lists, and 16 were semantic lures from other lists.

The 192 study items for each list were presented in a pseudorandom order. The pairs of parent items were arranged so that there were one to four intervening items between them. The test items were also presented in a pseudorandom order with no more than three items of the same type (old, lure, or new) presented in a row.

All participants in Experiment 2 were asked to complete the Shipley Vocabulary Test. Otherwise, the procedure for Experiment 2 was identical to that of Experiment 1.

Results

Tables 4 and 5 provide the mean proportions of each confidence rating for each item type for the older and younger adults. The number of "yes" responses in each condition was used to calculate the hit rate and false alarm rates for each participant. For the younger adults, there was a marginally significant difference in the average hit rates for the two conditions, with hit rates of 63% for the old items that were originally studied in a list context and 56% for the old items that were originally studied in sentences, $t(31) = 1.95$, $p = .06$. For the older adults, the average hit rate was 68% for the sentence condition and 66% for the out-of-context word condition; these two conditions were not significantly different, $t(31) = 0.58$. Comparing across the two age groups, we tested the hypothesis that the older adults would perform better than the younger adults for the sentence study condition but not for the word study condition. As predicted, the older adults had higher hit rates than the younger adults for words that were studied in sentences, $t(62) = 2.54$, $p = .01$ but not for words that were studied out-of-context, $t(62) = 1.03$, $p = .31$.

Figure 4 shows the false alarm rates for both groups of participants. A 2 × 2 × 2 ANOVA (study context × lure type × age) revealed a significant interaction between study context (word or sentence) and lure type (conjunction lure or semantic lure; $F(1, 62) = 7.37$, $p < .01$). There was no three-way interaction and

Table 4
Means and Standard Deviations for Proportions of Each Confidence Rating for Each Item Type for Older Adults in Experiment 2

		Confidence ratings			
Study context	Test item type	1 (Sure new)	2 (Unsure new)	3 (Unsure old)	4 (Sure old)
Sentence condition	Conjunction lures	0.33 (0.21)	0.32 (0.20)	0.20 (0.16)	0.15 (0.13)
	Semantic lures	0.37 (0.23)	0.27 (0.18)	0.15 (0.14)	0.21 (0.19)
	Old words	0.15 (0.10)	0.19 (0.14)	0.19 (0.14)	0.47 (0.17)
Word condition	Conjunction lures	0.32 (0.25)	0.30 (0.20)	0.23 (0.17)	0.15 (0.16)
	Semantic lures	0.40 (0.23)	0.27 (0.20)	0.17 (0.14)	0.16 (0.18)
	Old words	0.13 (0.11)	0.19 (0.17)	0.20 (0.14)	0.47 (0.23)
None	New items	0.43 (0.25)	0.31 (0.21)	0.16 (0.11)	0.10 (0.10)

no main effect of age. The results for both age groups replicate the effects of study context on susceptibility to lures that were observed in Experiment 1 and in Matzen and Benjamin (2009). As in the earlier experiments, participants were more susceptible to conjunction lures when the corresponding parent items were studied as out-of-context words as opposed to when the parent items were studied in sentence contexts. The opposite was true for semantic lures, with participants making more false alarms to semantic lures whose parents were studied in sentences than for those whose parent words were studied out-of-context.

The pattern of $\Delta\ d_a$ values across conditions was similar for both the younger and older adults and was consistent with the patterns observed in Experiment 1 and in Matzen and Benjamin (2009). The $\Delta\ d_a$ values for Experiment 2 are shown in Figure 5. Once again, the younger adults showed a bigger decrease in performance for the conjunction lures with parents that were studied as single words ($\Delta\ d_a = 0.33$) than for the conjunction lures with parents that were studied in sentences ($\Delta\ d_a = 0.22$). The older adults had the same pattern of performance ($\Delta\ d_a = 0.30$ for the single word condition; $\Delta\ d_a = 0.25$ for the sentence condition). Both younger and older adults had the opposite pattern of performance for the semantic lures. The younger adults had a bigger decrease in discrimination performance when the parent items of the semantic lures were studied in sentences ($\Delta\ d_a = 0.22$) than when the parent items were studied as single words ($\Delta\ d_a = 0.14$). The same was true for the older adults ($\Delta\ d_a = 0.26$ for the sentence condition; $\Delta\ d_a = 0.20$ for the single word condition). As with the false alarm rates, a 2 × 2 × 2 ANOVA showed that there was a significant interaction between lure type and study context for the $\Delta\ d_a$ values, $F(1, 62) = 7.14$, $p < .01$. In addition, there was a significant correlation between the participants' average $\Delta\ d_a$ values and their Shipley vocabulary scores ($r = .35$, $t(62) = 2.91$, $p < .01$). However, when the two groups of participants were examined separately, the correlation between average $\Delta\ d_a$ values and Shipley scores was significant for the younger adults, $r = .49$, $t(30) = 3.09$, $p < .01$, but not for the older adults, $r = .26$, $t(30) = 1.45$, $p = .08$.

We hypothesized that the older adults would have better discrimination performance than the young adults for the sentence study context but not for the word study context. The d_a results from Experiment 2, shown in Figure 6, support this hypothesis. For old–new discrimination, the older adults performed significantly better than the younger adults for words that were originally studied in sentences, $t(62) = 2.39$, $p = .02$ but there was not a significant difference between the two groups' performance for the items that were originally studied as single words, $t(62) = 0.17$. The same was true for the participants' ability to discriminate conjunction lures from old items: The older adults performed significantly better than the younger adults for lures with parents that were studied in sentences, $t(62) = 2.18$, $p = .03$, and the two groups' performance did not differ for the lures with parents that were studied as single words, $t(62) = 0.35$. For the semantic lures, the older adults had marginally better performance than the younger adults for the sentence condition, $t(62) = 1.75$, $p = .09$, and once again, the two groups' performance did not differ for the lures from the single word condition, $t(62) = 0.17$. The correlations between the participants' Shipley vocabulary scores and their average discrimination performance were not significant for either group: for younger adults, $r = .05$, $t(30) = 0.313$; for older adults, $r = .21$, $t(30) = 1.16$; for all participants, $r = .17$, $t(62) = 1.36$, $p = .09$.

Table 5
Means and Standard Deviations for Proportions of Each Confidence Rating for Each Item Type for Younger Adults in Experiment 2

		Confidence ratings			
Study context	Test item type	1 (Sure new)	2 (Unsure new)	3 (Unsure old)	4 (Sure old)
Sentence condition	Conjunction lures	0.33 (0.28)	0.36 (0.24)	0.19 (0.12)	0.13 (0.12)
	Semantic lures	0.37 (0.27)	0.33 (0.19)	0.14 (0.12)	0.16 (0.14)
	Old words	0.21 (0.22)	0.23 (0.16)	0.16 (0.12)	0.40 (0.18)
Word condition	Conjunction lures	0.29 (0.26)	0.35 (0.19)	0.21 (0.15)	0.15 (0.13)
	Semantic lures	0.38 (0.24)	0.35 (0.22)	0.14 (0.12)	0.12 (0.13)
	Old words	0.17 (0.18)	0.20 (0.16)	0.16 (0.12)	0.47 (0.21)
None	New items	0.41 (0.29)	0.37 (0.23)	0.14 (0.09)	0.08 (0.13)

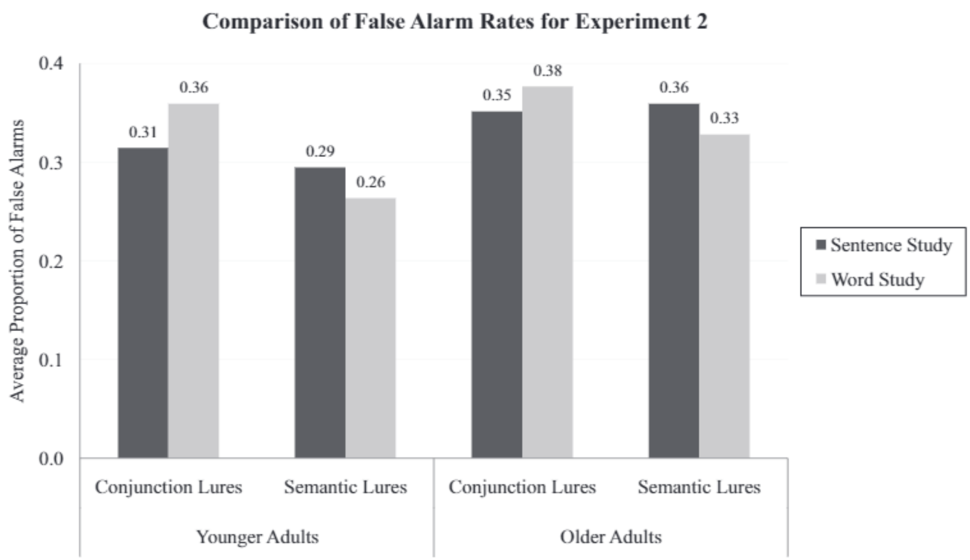

Figure 4. Comparison of false alarm rates for older and younger adults in Experiment 2.

Discussion

The results of Experiment 2 showed similar patterns of memory errors for both younger and older adults. The false alarm rates and the $\Delta\, d_a$ values both showed that there was a significant interaction between study context and lure type and that the same patterns emerged for younger and older adults. Experiment 2 also replicated the finding that the older adults generally had *better* discrimination performance than the younger adults in the sentence conditions. The younger and older adults had equivalent discrimination performance for the single word conditions, but the younger adults were less able to discriminate lures and new items from old items in the sentence study conditions. Although the older adults had higher vocabulary scores than the younger adults, the correlation between their overall discrimination performance and their vocabulary scores was not significant. The older adults who scored highest on the Shipley Vocabulary Test did not necessarily have high average $\Delta\, d_a$ values. This indicates that vocabulary alone was not the driving factor behind the older adults' superior performance in the sentence conditions.

General Discussion

The results of these experiments lead to two conclusions. First, Experiment 1 and Experiment 2 show that both older adults and

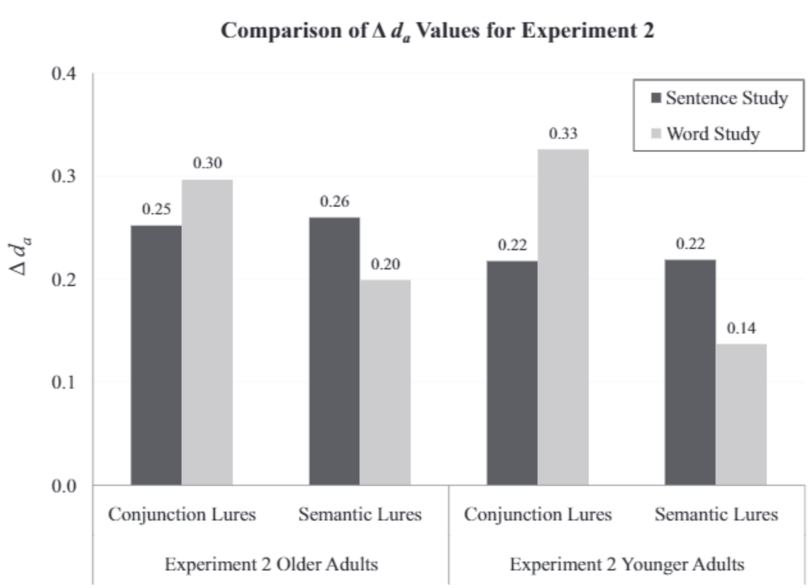

Figure 5. $\Delta\, d_a$ values for older and younger adults in Experiment 2. $\Delta\, d_a$ is the difference between the d_a value for old-new discrimination and the d_a value for old-lure discrimination for each condition.

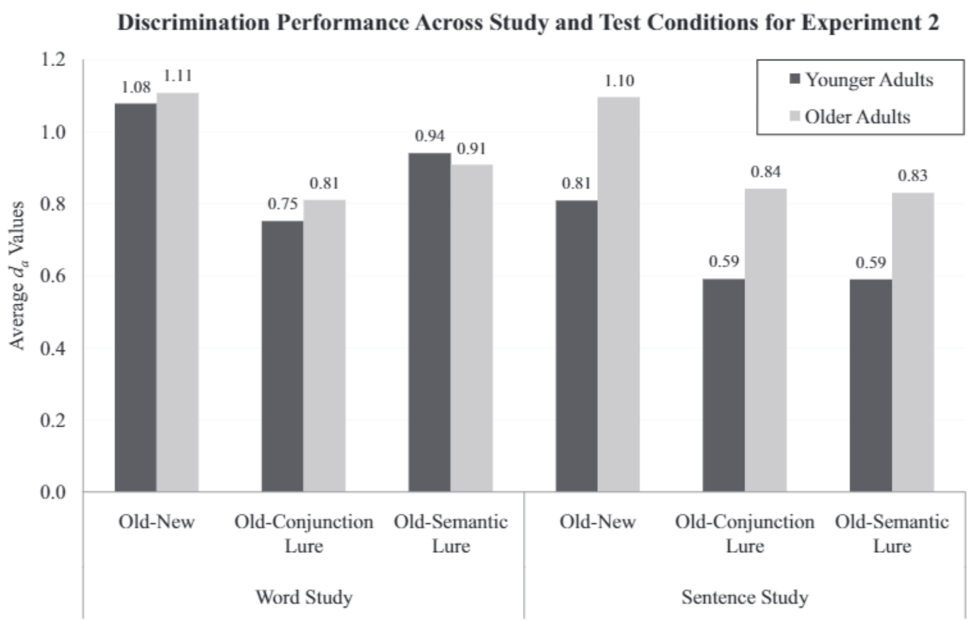

Figure 6. Comparison of discrimination performance (as measured by d_a) for the older and younger adults in Experiment 2.

younger adults are sensitive to changes in study context, leading to similar patterns of memory errors for the two groups. Second, and more surprisingly, the results of both experiments indicate that older adults exhibit better discrimination performance than younger adults for words that were previously presented in sentences. The advantage for the older adults remains even when evaluating recognition with semantic and conjunction lures that older adults are known to be especially susceptible to falsely remembering.

The Influence of Sentences on Encoding Biases

Matzen and Benjamin (2009) showed that, for younger adults, study context changed participants' susceptibility to different types of memory errors. When they studied sentences, the participants were less susceptible to conjunction lures but more susceptible to semantic lures. When they studied words in a list, the opposite was true, making participants more susceptible to conjunction lures and less susceptible to semantic lures. These patterns suggest that people encode words differently when they are presented with or without a larger meaningful context.

The effects of context on encoding may reflect participants' previous experience with the information that can be gleaned from different types of materials. In a sense, this sensitivity to study context can be thought of as a study "strategy." Although it is unlikely that participants consciously selected and alternated between two different study strategies when viewing single words or sentences, their prior experience with lists and sentences in daily life may have led the participants to shift the way in which they encoded the information in the two different contexts. This shift in encoding was apparent for both younger and older adults.

In addition to determining what sorts of information seem familiar at test, shifts in the nature of the information encoded during study also affect attempts at recollection. The process of recollection rejection, where participants are able to reject a lure by recalling its parent item, has been widely studied with respect to conjunction lures and other types of false memories (Brainerd & Reyna, 2002; Brainerd, Reyna, Wright, & Mojardin, 2003; Hintzman, Curran, & Oppy, 1992; Lampinen, Odegard, & Neuschatz, 2004; Matzen et al., 2011). When participants are successfully able to recall the originally studied words, they are also much more successful at rejecting the conjunction lures (Lampinen et al., 2004; Odegard & Lampinen, 2005; Odegard, Lampinen, & Toglia, 2005). Thus, factors that influence encoding can also influence recollection rejection by making the relevant information about the studied items more or less available. In the present study, placing compound words into sentences provides a richer context that can give participants additional cues to use for recollection rejection.

When the participants encountered words in a sentence context, they were likely to encode information about the gist of each sentence but relatively less information about the exact forms of the words they contained. Later, when semantic lures related to the sentences were tested, they were difficult to reject because they fit with the gist of a studied sentence but participants did not retain enough information about surface form to distinguish one semantic associate from another. The conjunction lures, which did not fit with the gist of any studied sentence, were easier to reject. When participants studied out-of-context words, there was little gist information available. This led participants to encode relatively more information about the surface forms of the words. This type of encoding made semantic lures easier to reject, because they did not overlap in surface form with the studied words. Yet it also led to more false alarms for conjunction lures, which share the surface forms of the studied words. Each type of encoding has advantages and disadvantages, as evidenced by the interaction between study context and susceptibility to the two types of lures.

Previous research on older adults' memory performance led us to consider two different predictions about older adults' performance on these tasks. Numerous studies have found that older adults are more susceptible to semantic errors than younger adults (cf. Balota et al., 1999; Norman & Schacter, 1997; Tun et al., 1998; Watson et al., 2001), that they are less able to use source information to avoid memory errors (Jones & Jacoby, 2005; Pierce et al., 2005), and that they are less able to use memory strategies in general (Anschutz et al., 1987; Jacoby, 1999; Kensinger & Schacter, 1999; Verhaeghen et al., 1992; Watson et al., 2004). These findings suggest that older adults would have high false alarm rates for semantic lures, regardless of study context. They also suggest that older adults' false alarms to conjunction lures would not decrease in the sentence study condition because they would be unable to use the information about the sentence contexts to reject the lures. From this perspective, one might not expect older adults to replicate the interaction between study context and lure type evident in younger adults.

However, other research indicates that older adults can use memory strategies effectively when those strategies are naturalistic (Castel, 2008; Dunlosky & Hertzog, 1998; Fraundorf et al., 2012; Rahhal, May, & Hasher, 2002; Rendell & Craik, 2000; Tentori et al, 2001; Tullis & Benjamin, 2012). As described above, the shifts in the type of information that was encoded from the two different study contexts can be thought of as being similar to a study strategy. Although this shift is probably not a deliberate choice on the part of the participants, their experiences with lists and sentences in daily life could lead them to approach the two types of contexts in slightly different ways. Older adults, with many more years of experience with reading and generally higher vocabulary and verbal knowledge than younger adults (cf. Kemper & Sumner, 2001; Park et al., 2002), may be able to adapt to the different study contexts even more effectively than younger adults.

Older Adults' Superior Discrimination Performance for Words in Sentences

If there is one story to be told about the effects of aging on episodic memory from the decades of research on the topic, it is a negative one. Older adults exhibit consistently poorer memory on an impressive variety of memory tests. This ubiquity is even more impressive given the generally higher levels of world and linguistic knowledge held by the older adults that form the population for many memory experiments. From this perspective, our finding that older adults exhibit superior memory to younger adults for words studied in sentences is quite noteworthy.

This result is perhaps best understood by remembering that the older adults' deficits often reveal a gradient of sorts, with tasks that require the most self-initiated processing revealing higher levels of age-related deficit than tasks that provide more environmental support (Craik, 1986). For example, older adults have much greater difficulty with free recall than recognition (Craik & McDowd, 1987).

The sentence study contexts used in these experiments can be conceptualized as another form of environmental support, though at the time of encoding rather than retrieval. Unlike word lists, which place a burden on the rememberer to bring effective organizational and encoding strategies to bear, sentences place a premium on skills that are honed through years of reading expertise. Recognition testing provides another form of support (Craik, 1986) and does not place much of a premium on the arbitrary and nonecological strategies that foster free recall of unrelated word lists. This can be seen in the classic result of Eagle and Leiter (1964), who showed that incidental encoding actually led to superior recognition performance relative to intentional encoding (though the opposite was true for free recall).

Older adults have extensive experience with reading sentences over their lifetimes, far more experience than the younger adults have had. Their experience and higher verbal knowledge may enable the older adults to make better use of the sentence contexts. In this particular experiment, higher verbal knowledge may have been particularly helpful in the conjunction lure conditions because many of the compound words used to form conjunction items are low frequency words and the older adults may have been more familiar with those words than the younger adults. This could explain why the older adults had a bigger advantage over the younger adults for conjunction lures in the sentence condition than for semantic lures in the sentence condition.[1]

It is also possible that the older adults engaged in more elaborative processing while reading the sentences. Past research has shown that background knowledge is correlated with recollection in memory for texts (Long & Prat, 2002; Long, Prat, Johns, Morris, & Jonathan, 2008). The scenarios described in the sentences may remind the participants of their own prior experiences (anecdotally, some of the older participants reported this after completing the experiment), and the older adults have more experiences to draw upon. Elaboration or relating the sentences to their own experiences could aid the older adults in recollection of the old items and in recollection rejection of the lures by helping them to remember specific details about the sentences. In general, stimuli that older adults show an interest in or are instructed to attend to tend to reveal smaller age-related deficits, if any at all (Benjamin et al., 2012; May, Rahhal, Berry, & Leighton, 2005; Rahhal et al., 2002).

In summary, the results of these experiments show that older adults' memory performance can be affected by study context, leading to patterns of memory errors that are very similar to those seen for younger adults. Sentence processing is a natural task with which older adults have a great deal of experience. The naturalistic nature of the constraints provided by the two different study contexts may explain why the older adults' and younger adults' patterns of memory errors are so similar. Additionally, the older adults exhibited better recognition than the younger adults for information that was studied in sentences. Though memory in older adults is often poorer than in younger adults, conditions that play to the strengths of the elderly—such as experience with reading and greater knowledge—and minimally to their weaknesses—such as self-initiated processing—have the potential to reveal a

[1] It is possible that participants were sensitive to the high number of compound words on the study and test lists, making them more likely to respond "new" to noncompound words in general. Despite the fully counterbalanced design of Experiment 2, the conjunction lures were the only test items that were always compound words. Thus, differential sensitivity between the two age groups to the preponderance of compound words could affect the conjunction lures (exclusively compound words) differently than the other conditions, which included some noncompound words. However, this would not affect our overall conclusions.

more complex, and perhaps more positive view, of the effects of aging on memory.

References

Alba, J., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93,* 203–231. doi:10.1037/0033-2909.93.2.203

Anschutz, L., Camp, C. J., Markley, R. P., & Kramer, J. J. (1987). Remembering mnemonics: A three-year follow-up on the effects of mnemonics training in elderly adults. *Experimental Aging Research, 13,* 141–143. doi:10.1080/03610738708259315

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

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). *The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords*. St. Louis, MO: Washington University. Retrieved from http://elexicon.wustl.edu/

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11,* 267–273. doi:10.1111/1467-9280.00254

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

Benjamin, A. S. (2005). Recognition memory and introspective remember/know judgments: Evidence for the influence of distractor plausibility on "remembering" and a caution about purportedly nonparametric measures. *Memory & Cognition, 33,* 261–269. doi:10.3758/BF03195315

Benjamin, A. S. (2010). Representational explanations of "process" dissociations in recognition: The DRYAD theory of aging and memory judgments. *Psychological Review, 117,* 1055–1079. doi:10.1037/a0020810

Benjamin, A. S., Diaz, M., Matzen, L. E., & Johnson, B. (2012). Tests of the DRYAD theory of the age-related deficit in memory for context: Not about context, and not about aging. *Psychology and Aging, 27,* 418–428.

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review, 116,* 84–115. doi:10.1037/a0014351

Bissig, D., & Lustig, C. (2007). Who benefits from memory training? *Psychological Science, 18,* 720–726. doi:10.1111/j.1467-9280.2007.01966.x

Bock, J. K., & Brewer, W. F. (1974). Reconstructive recall in sentences with alternative surface structures. *Journal of Experimental Psychology, 103,* 837–843. doi:10.1037/h0037391

Brainerd, C. J., & Reyna, V. F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology, 38,* 156–172. doi:10.1037/0012-1649.38.1.156

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110,* 762–784. doi:10.1037/0033-295X.110.4.762

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331–350. doi:10.1016/0010-0285(71)90019-3

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition, 5,* 673–678. doi:10.3758/BF03197414

Castel, A. D. (2008). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. In A. S. Benjamin & B. H. Ross (Eds.), *The psychology of learning and motivation* (Vol. 48, pp. 225–270). San Diego: CA: Academic Press.

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition, 30,* 1078–1085. doi:10.3758/BF03194325

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409–422). Amsterdam, The Netherlands: Elsevier.

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. doi: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. doi:10.1037/h0046671

Dunlosky, J., & Hertzog, C. (1998). Training program to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah, NJ: Erlbaum.

Eagle, M., & Leiter, E. (1964). Recall and recognition in intentional and incidental learning. *Journal of Experimental Psychology, 68,* 58–63. doi:10.1037/h0044655

Egan, J. P. (1958). *Recognition memory and the operating characteristic (Tech. Note AFCRC-TN-58_51)*. Bloomington, IN: Indiana University, Hearing and Communication Laboratory.

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source- relevant cues. *Psychology and Aging, 7,* 443–452. doi:10.1037/0882-7974.7.3.443

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2012). The effects of age on the strategic use of pitch accents in memory for discourse: A processing-resource account. *Psychology and Aging, 27,* 88–98. doi:10.1037/a0024138

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Halamish, V., McGillivray, S., & Castel, A. D. (2011). Monitoring one's own forgetting in younger and older adults. *Psychology and Aging, 26,* 631–635. doi:10.1037/a0022852

Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General, 127,* 251–268. doi:10.1037/0096-3445.127.3.251

Hertzog, C., & Dunlosky, J. (2005). Aging, metacognition, and cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 45, pp. 215–251). San Diego, CA: Academic Press.

Hintzman, D. L., Curran, T., & Oppy, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18,* 667–680. doi:10.1037/0278-7393.18.4.667

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 3–22. doi:10.1037/0278-7393.25.1.3

Johnson, M. K., Bransford, J. D., & Solomon, S. K. (1973). Memory for tacit implications in sentences. *Journal of Experimental Psychology, 98,* 203–205. doi:10.1037/h0034290

Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language, 45,* 82–102. doi:10.1006/jmla.2000.2761

Jones, T. C., & Jacoby, L. L. (2005). Conjunction errors in recognition memory: Modality-free errors for older adults but not for young adults. *Acta Psychologica, 120,* 55–73. doi:10.1016/j.actpsy.2005.03.003

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16,* 312–322. doi:10.1037/0882-7974.16.2.312

Kensinger, E. A., & Schacter, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology, 16,* 399–415. doi:10.1080/026432999380852

Kroll, N. E. A., Knight, R. T., Metcalfe, J., Wolf, E. S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language, 35,* 176–196. doi:10.1006/jmla.1996.0010

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lampinen, J. M., Odegard, T. N., & Neuschatz, J. S. (2004). Robust recollection rejection in the memory conjunction paradigm. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 30, 332–342. doi:10.1037/0278-7393.30.2.332

Light, L. L. (1996). Memory and aging. In E. K. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 443–490). San Diego, CA: Academic Press. doi:10.1016/B978-012102570-0/50015-X

Long, D. L., & Prat, C. S. (2002). Memory for Star Trek: The role of prior knowledge in recognition revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1073–1082. doi:10.1037/0278-7393.28.6.1073

Long, D. L., Prat, C. S., Johns, C., Morris, P., & Jonathan, E. (2008). The importance of knowledge in vivid text memory: An individual-differences investigation of recollection and familiarity. *Psychonomic Bulletin & Review, 15,* 604–609. doi:10.3758/PBR.15.3.604

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Matzen, L. E., & Benjamin, A. S. (2009). Remembering words not presented in sentences: How study context changes patterns of false memories. *Memory & Cognition, 37,* 52–64. doi:10.3758/MC.37.1.52

Matzen, L. E., Taylor, E. T., & Benjamin, A. S. (2011). Contributions of familiarity and recollection rejection to recognition: Evidence from the time course of false recognition for semantic and conjunction lures. *Memory, 19,* 1–16. doi:10.1080/09658211.2010.530271

May, C. P., Rahhal, T., Berry, E. M., & Leighton, E. A. (2005). Aging, source memory, and emotion. *Psychology and Aging, 20,* 571–578. doi:10.1037/0882-7974.20.4.571

McIntyre, J. S., & Craik, F. I. M. (1987). Age differences in memory for item and source information. *Canadian Journal of Psychology, 41,* 175–192. doi:10.1037/h0084154

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.

Odegard, T. N., & Lampinen, J. M. (2005). Recollection rejection: Gist cuing of verbatim memory. *Memory & Cognition, 33,* 1422–1430. doi:10.3758/BF03193375

Odegard, T. N., Lampinen, J. M., & Toglia, M. P. (2005). Meaning's moderating effect on recollection rejection. *Journal of Memory and Language, 53,* 416–429. doi:10.1016/j.jml.2005.04.004

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17,* 299–320. doi:10.1037/0882-7974.17.2.299

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory, 4,* 171–212.

Pierce, B. H., Sullivan, A. L., Schacter, D. L., & Budson, A. E. (2005). Comparing source-based and gist-based false recognition in aging and Alzheimer's disease. *Neuropsychology, 19,* 411–419. doi:10.1037/0894-4105.19.4.411

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29,* 633–654. doi:10.1016/0749-596X(90)90042-X

Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science, 13,* 101–105. doi:10.1111/1467-9280.00419

Reinitz, M. T., Lammers, W. J., & Cochran, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition, 20,* 1–11. doi:10.3758/BF03208247

Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age related differences in prospective memory. *Applied Cognitive Psychology, 14,* S43–S62. doi:10.1002/acp.770

Roediger, H. L., & McDermott, K. (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

Roediger, H. L., Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 3–39). Hillsdale, NJ: Erlbaum.

Rubin, S. R., Van Petten, C., Glisky, E. L., & Newberg, W. M. (1999). Memory conjunction errors in younger and older adults: Event-related potential and neuropsychological data. *Cognitive Neuropsychology, 16,* 459–488. doi:10.1080/026432999380889

Schacter, D. L., Kaszniak, A. W., Kihlstrom, J. F., & Valdiserri, M. (1991). The relation between source memory and aging. *Psychology and Aging, 6,* 559–568. doi:10.1037/0882-7974.6.4.559

Schacter, D. L., Koustaal, W., Johnson, M. K., Gross, M. S., & Angell, K. A. (1997). False recollection induced by photographs: A comparison of older and younger adults. *Psychology and Aging, 12,* 203–215. doi:10.1037/0882-7974.12.2.203

Shipley, W. C. (1946). *Institute of living scale*. Los Angeles, CA: Western Psychological Services.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10,* 527–539. doi:10.1037/0882-7974.10.4.527

Swets, J. A. (1986). Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. *Psychological Bulletin, 99,* 181–198. doi:10.1037/0033-2909.99.2.181

Tentori, K., Osherson, D., Hasher, L., & May, C. (2001). Wisdom and aging: Irrational preferences in college students but not older adults. *Cognition, 81,* B87–B96. doi:10.1016/S0010-0277(01)00137-8

Tullis, J. G., & Benjamin, A. S. (2012). The effectiveness of updating metacognitive knowledge in the elderly: Evidence from metamnemonic judgments of word frequency. *Psychology and Aging, 27,* 939.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373. doi:10.1037/h0020071

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

Underwood, B. J., Kapelak, S. M., & Malmi, R. A. (1976). Integration of discrete verbal units in recognition memory. *Journal of Experimental Psychology: Human Learning & Memory, 2,* 293–300. doi:10.1037/0278-7393.2.3.293

Underwood, B. J., & Zimmerman, J. (1973). The syllable as a source of error in multisyllable word recognition. *Journal of Verbal Learning & Verbal Behavior, 12,* 701–706. doi:10.1016/S0022-5371(73)80050-7

Verhaeghen, P., Marceon, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging, 7,* 242–251. doi:10.1037/0882-7974.7.2.242

Watson, J. M., Balota, D. A., & Sergent-Marshall, S. D. (2001). Semantic, phonological, and hybrid veridical and false memories in healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology, 15,* 254–268. doi:10.1037/0894-4105.15.2.254

Watson, J. M., McDermott, K. B., & Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger–McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition, 32,* 135–141. doi:10.3758/BF03195826

Received May 9, 2012
Revision received March 26, 2013
Accepted April 1, 2013 ■