Aging, spaced retrieval, and inflexible memory performance

Anthony J. Bishara

Indiana University

Larry L. Jacoby

Washington University

Reference:

Bishara, A. J., & Jacoby, L. L. (2008). Aging, spaced retrieval, and inflexible memory performance. *Psychonomic Bulletin & Review, 15*, 52-57.

Abstract

Spaced retrieval is a memory-training technique whereby information is tested at progressively longer delays. Two experiments were conducted to examine effects of spaced retrieval on controlled recollection and automatic influences of memory. In Experiment 1, word-pairs were read once, read three times, or read once and retrieved twice by young and older adults. Retrieval practice improved performance on a later test for both age groups. Experiment 2 was arranged such that recollection opposed automatic influences of retrieval practice. Retrieval practice increased intrusions on a later test only for older adults. Results suggest that, because of a deficit in recollection, older adults were less able to oppose the automatic influence of spaced retrieval, and so exhibited less flexible memory performance.

Aging, spaced retrieval, and inflexible memory performance

A great deal of research has shown that testing improves memory, sometimes even more than additional studying does (e.g., Carrier & Pashler, 1992; Jacoby & Bartz, 1972; Roediger & Karpicke, 2006). Repeated testing with progressively increasing delays between tests, termed *expanded* or *spaced retrieval*, can greatly enhance memory performance (Landauer & Bjork, 1978). Spaced retrieval has been used to enhance memory performance in various populations, ranging from Alzheimer's patients (Camp, Foss, O'Hanlon, & Stevens, 1996; Schacter, Rich, & Stampp, 1985) to healthy older adults (Logan & Balota, in press). Despite these successful applications, spaced retrieval may come with the cost of leading to inflexible behavior in populations with deficits in controlled memory functioning, including aging populations.

To illustrate this problem, consider a study in which spaced retrieval was used to train participants with Alzheimer's disease to open a box each time an alarm sounded (Bird & Kinsella, 1996). Participants did learn to open the box. However, the learning was inflexible and not restricted to the alarm being sounded. For example, one participant inappropriately shouted "open the box" the next time she saw the experimenter.

Anecdotes aside, there is currently no published empirical work examining whether spaced retrieval leads to inflexible performance. One published study has examined spaced retrieval in terms of a related issue, a comparison of explicit and implicit memory (Cherry, Simmons, & Camp, 1999). Unfortunately, spaced retrieval's effects on explicit and implicit memory were ambiguous because the small sample size prevented statistical hypothesis testing. Closely related to spaced retrieval is the technique of vanishing cues, whereby memory tests are made more difficult by progressively removing cues rather than progressively increasing test delays (Glisky, Schacter, & Tulving, 1986). Learning from vanishing cues has been shown to sometimes lead to inflexible behavior (Kapur, Glisky, & Wilson, 2004).

The current article examines inflexible behavior by comparing spaced retrieval to spaced studying. Experiment 1 shows that, compared to spaced studying, spaced retrieval can improve memory performance for both young and older adults. Results from Experiment 2 suggest that spaced retrieval enhanced both controlled recollection and more automatic influences of memory for young adults. However, recollection is more susceptible to the effects of aging (e.g., Jennings & Jacoby, 1993, 1997) and so for older adults, spaced-retrieval practice might primarily increase automatic influences, thereby leading to inflexible behavior.

Experiment 1

Memory performance in a *Retrieved* condition was compared with that produced by studying word pairs simply by reading them once or three times (*Read 1x* and *Read 3x* conditions). For the Retrieved condition, we followed Carrier and Pashler (1992) by providing additional exposure to responses when retrieval attempts failed. Doing so is important because otherwise failures to retrieve items on initial tests might hide advantages of retrieval over repeated reading. The Retrieved condition involved three trials. On the first trial, associatively related pairs of words (e.g., knee-bone) were presented for study. On the second and third trials, the cue word of the pair was accompanied by a fragment of the response (knee-b_n_), and participants were instructed to respond with the whole word pair. For the final second of the two trials for which retrieval was attempted, the cue was presented with its intact response (knee-bone) regardless of whether or not retrieval was successful.

An expanded retrieval schedule was used such that there was more intervening time between the 2nd and 3rd trials than between the 1st and 2nd trials. This expanding schedule was also used for the Read 3x condition so that it would be comparable. In a later test phase, memory was measured for all three conditions by presenting the cue word along with a fragment of its response (e.g., knee-b_n_). We expected memory performance to be better in the Retrieved condition than the Read 3x condition, which in turn was expected to be better than the Read 1x condition.

Method

Participants

Table 1 shows demographics of participants. The young adult group consisted of Washington University undergraduates who participated in exchange for course credit or $10. The older adult group was recruited from the Washington University Psychology Department older adult participant pool. Older adults were volunteers from the St. Louis community and participated in exchange for $10.

Design and Materials

The two age groups were crossed with three within-subject conditions: Read 1x, Read 3x, and Retrieved.

A pool of 94 word triplets was selected from norms reported by Jacoby (1996). Each triplet included one cue word (e.g. knee) and two associatively related responses (bone, bend) that would complete a word fragment (b_n_). Of these triplets, 84 were used as critical items and divided into sets of 28 for the 3 conditions. Assignment of sets to conditions was counterbalanced across participants. The remaining 10 triplets were used as fillers. To reduce serial position effects, 2 fillers were presented at the beginning and 2 at the end of the study list. Six filler items were distributed throughout the list in order to obtain the desired repetition schedules.

A distracter list consisted of 84 new, once-presented word pairs (e.g. zoo-cage) that were chosen from association norms (Jacoby, 1996; Nelson, McEvoy, & Schreiber, 1998). The distracter list intervened between presentation of the study list and its test, and it was meant to reduce recall so as to avoid ceiling effects.

The test list consisted of 84 critical cue words presented with a word fragment (e.g. knee-b_n_). The fragments were the same as those presented in the Retrieved condition. Order of test presentation was random with the restriction that no more than two items from the same condition could be presented in succession.

Procedure

For the study phase, participants were instructed to read aloud and study word-pairs for a later memory test. They were told that some pairs would be repeated with letters missing from the second word (Retrieved condition). For these pairs, they were to recall the studied word that completed the fragment and say the whole pair out loud as soon as possible. In the Retrieved condition, the complete word pair appeared for 4 seconds for the first presentation. On the second and third presentations, the cue word and word fragment (e.g. knee-b_n_) appeared for a maximum of three seconds. As soon as the participant responded or after 3 seconds had elapsed, the word fragment was replaced by the intact pair. The intact pair remained on the screen for 1 second. Participants read the intact pair aloud if they had failed to respond in the previous 3 seconds. Read 1x and Read 3x word pairs appeared in complete form for 4 seconds on each presentation. For Read 3x and Retrieved items, the first presentation was separated from the second by 1-2 intervening pairs, and the second presentation was separated from the third by 5-6 intervening pairs.

For the distracter phase, participants were instructed to read aloud and to remember the word pairs for later. Pairs were presented for 2 seconds each. In actuality, there was no test for these items.

For the final test phase, participants were tested for their memory of the original study list. Each test trial consisted of a cue word and word fragment (knee-b_n_) appearing on the screen until the participant responded or for a maximum of 10 seconds. Participants were instructed to recall the response and say the whole word pair within that time. They were instructed to give a response even if they had to guess. A blank screen appeared between each trial for 500 ms. After the memory test, participants took a computer version of the Shipley Vocabulary Subtest.

Results and Discussion

While lags were chosen to minimize errors during the study phase, errors still occurred occasionally. The probability of producing at least 1 error out of 2 opportunities to do so during retrieval practice of an item was lower for young adults (*M* = .10, *SD* = .07) than older adults (*M* = .23, *SD* = .13), $t(34) = 3.84$, $p < .001$. These errors may strengthen incorrect responses, and so the main analysis of the final memory test is conditionalized on errorless retrieval during the study phase. For both experiments, unconditionalized data are reported in parentheses in tables. Statistical analysis of unconditionalized data is only reported when it changes significance status, and in most cases it did not.

Table 2 shows the probability of correct responding on the final test across Item Type and Age. Performance was better for young adults (*M* = .83) than older adults (*M* = .69). When a 2 (Age:Young & Older) X 3 (Item Type: Read 1x, Read 3x, & Retrieved) mixed ANOVA was performed, there was a significant main effect of Age, $F(1,34) = 24.55$, $p < .0001$, $\eta_p^2 = .42$.

Performance was lowest in the Read 1x condition (*M* = .66), higher in the Read 3x condition (.75), and highest in the Retrieved condition (.92). This was confirmed by a significant main effect of Item Type, $F(2,68) = 94.03$, $p < .00001$, $\eta_p^2 = .73$. Post-hoc paired t-tests showed that all three conditions were significantly different from one another for both age groups, all *p*s < .01. The interaction of Age and Item Type was not significant, $F(2,68) = 1.53$, $p > .10$, though it is possible that ceiling effects in the Retrieved condition obscured an interaction.

Overall, reading a pair once and then retrieving its response twice produced better memory performance than did studying the pair three times. Although retrieval practice enhanced memory performance, it did not remove the memory disadvantage of older adults.

Experiment 2

Retrieval practice might have the automatic influence of increasing the accessibility of responses, and also have a separate effect on controlled recollection. Experiment 2 examined age differences in these two effects of retrieval practice by using an opposition procedure similar to that used by Jacoby (1999) to examine age differences in the effects of repetition.

In Jacoby's experiments, young and older adults *read* a list of words, with each word being read once, twice, or three times. Next, they *heard* a second list that they were told to remember. At test, participants were instructed to identify words that they had heard earlier, and were warned that the test list would include earlier-read words. They were correctly informed that none of the earlier-read words had been presented in the heard list. Repeatedly reading a word was expected to increase its familiarity (an automatic influence of memory) and thus increase the probability that it would be mistakenly judged as heard. However, controlled recollection of having read a word would oppose its familiarity, allowing it to be excluded.

The results showed a significant age-by-repetition interaction, suggesting that repetition had two effects. For older adults, repeated reading increased the probability of words being mistakenly accepted as earlier heard, showing the effect of repetition on familiarity. In contrast, young adults were better able to use recollection to oppose familiarity so as to avoid accepting earlier-read words. The age-by-repetition interaction suggests that repetition enhanced automatic influences of memory, and at least for young adults, also increased controlled recollection that was used to successfully oppose the increased automatic influence.

In the opposition procedure used in Experiment 2, increased automaticity produced by retrieval practice would serve as a source of interference. The procedure is outlined in Table 3. The target responses for cue words (e.g., knee-bone) were presented for study in Phase 2, and participants were instructed to recall these for the final test. The first phase of Experiment 2 was the same as that of Experiment 1 except that the responses that were presented and retrieved (e.g., knee-bend) served as competitors for the target responses presented in Phase 2. Increased automaticity produced by retrieval practice would result in increased interference from Phase 1 responses (e.g., “bend” intruding when “bone” was intended to be recalled).

Experiment 2 differed from experiments by Jacoby (1999) in that it examined cued-recall rather than recognition-memory performance, and also differed in several other procedural details. Consequently, we did not expect the pattern of results to be identical to that observed by Jacoby (1999). However, we did expect older adults, as compared to young adults, to be less able to use recollection to oppose automatic influences of memory resulting from retrieval practice. As support for that prediction, we expected that retrieval practice would increase proactive interference, increasing intrusions of the practiced response more so for older adults than for young adults. That is, we predicted an interaction between retrieval practice (as compared to repeated reading) and age.

Experiment 2 included a guessing condition (see Table 3) that was identical to the retrieved condition except that the target response was not studied in Phase 2. This condition was meant to explore the possibility that the retrieval practice effects would be general, such that young adults would show reduced intrusions relative to older adults even when the target response had not been presented in the context of the experiment. Alternatively, the benefits of retrieval might not generalize to this condition because of failures to generate a substitute response. For example, a participant might recollect that “rabbit-hole” was studied in Phase 1, but then fail to generate “hare” as a substitute. In this case, the intrusion response “hole” might still be given.

One group of young adults and two groups of older adults participated in Experiment 2. The *Unmatched Older Adults* group followed the same procedure as young adults but showed worse baseline performance, which raised concerns about interpreting results from the Retrieved condition. Consequently, a second group of older adults (*Matched Older Adults*) was given extra study time in Phase 2 so as to match their baseline performance to that of young adults. We expected the age-by-retrieval interaction to occur even when baseline performance was matched.

Method

Participants

There were three groups: Young Adults, Unmatched Older Adults, and Matched Older Adults (see Table 1). Though the Matched group was recruited after the Unmatched group, the two older groups were recruited from the same pool and did not differ from each other in terms of age, vocabulary, education, or health, all *p*s > .10.

Design, Material, and Procedure

Only changes relative to Experiment 1 are described here. Five experimental conditions (see Table 3) were each represented by 20 word triplets. Assignment of word triplets to conditions was counterbalanced across participants, as was assignment of response words to be intrusions or targets.

For the final test, participants were instructed to respond with the word presented in Phase 2. Participants were questioned after the final test to assure that they understood the instructions. Three older adults were excluded and replaced because they were unclear about which phase they were to recall from.

The Read 1x, Read 3x, and Retrieved conditions were the same as in Experiment 1 except that the intrusion response, rather than the target, was presented in Phase 1. In the *Guessing* condition, retrieval practice of the competing response was provided in Phase 1, but neither the cue nor the target response was presented in Phase 2. Participants had to guess the correct response on the final test (e.g., “hare”; see Table 3). In the *Baseline* condition, the relevant pair was presented in Phase 2 only, and neither the cue nor the intrusion response was presented in Phase 1.

All pairs presented in Phase 2 were presented once and at a rate of 2 seconds for Young and Unmatched Older Adults, and 6 seconds for Matched Older Adults.

Results and Discussion

The probability of producing at least 1 error for an item during retrieval practice in Phase 1 was lower for Young adults (M = .09) than for Unmatched and Matched Older adults (Ms = .24 & .27, respectively), $F(2,57) = 15.28$, $p < .00001$, $\eta^2 = .35$.

Correct Recall

Table 4 shows the probability of correct responses (from Phase 2) and intrusions (from Phase 1). As indicated by symbols in the table, matching was successful in that t-tests showed no difference between Matched Older adults and Young adults on Baseline correct recall. In contrast, Unmatched Older adults performed worse than the young.

Importantly, aging was associated with less correct recall in the Retrieved condition than the Read 3x condition, and this was true for both older adult groups. An ANOVA was performed on correct responses in all 5 conditions X all 3 groups so as to insure that the overall interaction was significant, which it was, with the Greenhouse-Geisser correction for violated sphericity, $F(6.8,194.8) = 3.34$, $p < .01$, $\eta_p^2 = .10$. Next, a more restricted 2 X 2 ANOVA compared Young and Unmatched Older Adults in the Read 3x and Retrieved conditions to reveal a significant interaction, $F(1,38) = 6.47$, $p < .05$, $\eta_p^2 = .15$. Finally, a 2 X 2 ANOVA also revealed a significant interaction when comparing Young and Matched Older Adults in Read 3x and Retrieved conditions, $F(1,38) = 5.26$, $p < .05$, $\eta_p^2 = .12$.

Intrusions

Analyses of intrusions suggest that the poor performance of older adults in the retrieved condition was due to increased intrusions of the earlier retrieved response. First, the overall 5 X 3 ANOVA revealed a significant interaction, with the Greenhouse-Geisser correction $F(6.8,194.7) = 2.86$, $p < .01$, $\eta_p^2 = .09$. Next, a 2 X 2 ANOVA showed a significant interaction when comparing Young and Unmatched Older Adults in the Read 3x and Retrieved conditions, $F(1,38) = 7.65$, $p < .01$, $\eta_p^2 = .17$. Finally, a 2 X 2 ANOVA also showed a significant interaction when comparing Young and Matched Older Adults in Read 3x and Retrieved conditions, $F(1,38) = 4.89$, $p < .05$, $\eta_p^2 = .11$.[1]

Similar to what was found for study repetition in recognition memory (Jacoby, 1999), retrieval practice had separate effects on controlled and automatic influences in cued recall. Retrieval practice had the automatic influence of increasing the accessibility of the practiced response, which could serve as a source of interference. For young adults, retrieval practice also enhanced controlled recollection, which was used to successfully oppose this increased automatic influence. However, because of older adults' lessened ability to recollect, the automatic, interfering influence of retrieval practice was largely unopposed. The age difference in interfering effects of retrieval practice was observed even when older adults were equated to young adults in a baseline condition.

There was a trend in the means suggesting that young adults had fewer intrusions in the Guessing condition, too. However, the effect of group was not significant there, as shown by a one-way ANOVA, $F(2,57) = 1.67$, $p = .20$. It is possible that, even when recollection succeeded, participants often gave the intrusion response because it was difficult to generate a suitable substitute when none had been presented in the context of the experiment. Indeed, it was rare for suitable substitutes to be reported in the Guessing condition (this is labeled "mean correct recall" in Table 4). Age differences in retrieval effects might be most apparent when at least two appropriate responses are readily available.

Neither repetition nor retrieval practice caused a significant change in intrusions in young adults' cued-recall performance. In contrast, Jacoby (1999) found that repetition caused a decrease in false alarms in recognition-memory performance. Different results are unsurprising given the different measures of memory and other procedural differences. Repetition or retrieval would be expected to cause a significant decrease in interference only if the effects on recollection were larger than those on automatic influences of memory. More research is needed to uncover the factors that are responsible for the balance of the two effects in both repetition and spaced retrieval.

Importantly, the pattern of results across experiments does suggest that spaced retrieval had two effects for young adults rather than no effect. In Experiment 1, spaced retrieval did enhance memory performance for young adults. Given this result, it would be difficult to explain the age-by-retrieval interaction in Experiment 2 without postulating opposing processes along with age differences in recollection.

General Discussion

Results of Experiment 1 showed that spaced-retrieval practice enhanced later memory performance more than did additional study. Results of Experiment 2 indicate that spaced retrieval had the automatic effect of increasing the accessibility of practiced responses, and the second effect of enhancing controlled recollection. The balance of these two effects places important limitations on the utility of retrieval practice as a means of improving memory. When a person's ability to recollect is impaired, the facilitation gained by increasing the accessibility of correct responses (Experiment 1) is offset by the cost when those highly accessible responses are incorrect ones (Experiment 2).

We have argued that retrieval enhances two processes, though without explaining why retrieval would enhance either. Carrier and Pashler (1992) suggested that retrieval benefits can potentially be explained by network models that learn through error-reduction mechanisms. In such models, an external presentation of a response (such as in the Read 3x condition) can spoil the network's ability to generate its own noisy representation and then learn from it. With this in mind, the most promising network models of retrieval effects on memory might be models that include error-correction learning mechanisms within a dual-process or dual-system framework (e.g., O'Reilly & Rudy, 2000).

Though the current results were framed in terms of a recollection impairment in older adults, it should be noted that the results are also consistent with age-related impairments in other varieties of controlled processing, such as source memory (Benjamin & Craik, 2001) or goal neglect (Jacoby, Bishara, Hessels, & Toth, 2005). Additional research is required to choose among more fine-grained distinctions regarding controlled processing. Overall, the current results suggest that older adults and perhaps other populations with controlled memory impairments might suffer from inflexible behavior following spaced retrieval practice. This should be kept in mind when spaced retrieval is being considered as a tool for memory rehabilitation.

References

Benjamin, A. S., & Craik, F. I. M. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition, 29*(5), 691-697.

Bird, M., & Kinsella, G. (1996). Long-term cued recall of tasks in senile dementia. *Psychology and Aging, 11*, 45-56.

Camp, C. J., Foss, J. W., O'Hanlon, A. M., & Stevens, A. B. (1996). Memory interventions for persons with dementia. *Applied Cognitive Psychology, 10*(3), 193-210.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory and Cognition, 20*(6), 633-642.

Cherry, K. E., Simmons, S. S., & Camp, C. J. (1999). Spaced retrieval enhances memory in older adults with probable Alzheimer's disease. *Journal of Clinical Geropsychology, 5*, 159-175.

Glisky, E. L., Schacter, D. L., & Tulving, E. (1986). Learning and retention of computer-related vocabulary in memory-impaired patients: Method of vanishing cues. *Journal of Clinical and Experimental Neuropsychology, 8*(3), 292-312.

Jacoby, L. L. (1996). Dissociating automatic and consciously-controlled effects of study/test compatibility. *Journal of Memory and Language, 35*, 32-52.

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.

Jacoby, L. L., & Bartz, W. H. (1972). Rehearsal and transfer to LTM. *Journal of Verbal Learning & Verbal Behavior, 11*(5), 561-565.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General, 134*, 131-148.

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging, 8*(2), 283-293.

Jennings, J. M., & Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology and Aging, 12*(2), 352-361.

Kapur, N., Glisky, E. L., & Wilson, B. A. (2004). Technological memory aids for people with memory deficits. *Neuropsychological Rehabilitation, 14*(1-2), 41-60.

Logan, J. M., & Balota, D. A. (in press). Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, & Cognition.*

Landauer, T. K., & Bjork, R. A. (1978). Optimal rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Harris, & R. N. Sykes (Eds.), *Practical aspects of memory*. New York: Academic Press.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. *Journal of Memory and Language, 35*, 32-52. See http://w3.usf.edu/FreeAssociation/.

O'Reilly, R. C., & Rudy, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus, 10*(4), 389-397.

Roediger, H.L., & Karpicke, J.D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249-255.

Schacter, D. L., Rich, S. A., & Stampp, M. S. (1985). Remediation of memory disorders: Experimental evaluation of the spaced-retrieval technique. *Journal of Clinical and Experimental Neuropsychology*, *7*(1), 79-96.

Author Note

Anthony J. Bishara, Department of Psychological & Brain Sciences, Indiana University, and Larry L. Jacoby, Department of Psychology, Washington University. We thank Danielle Hirschfeld, Carole Jacoby, Alex Konkel, Meredith Reuter, Chad Rogers, Michael Schachter, and Aaron Shumaker for helping with data collection, data coding, and other aspects of this research. This research was supported by a grant from the National Institute on Aging AG13845.

Correspondence should be sent to Anthony J. Bishara, Indiana University, Department of Psychological & Brain Sciences, 1101 East Tenth Street, Bloomington, IN 47405. Electronic mail may be sent to abishara@indiana.edu.

Footnotes

[1] In Experiment 2, the 2 (Young vs. Matched Older Adults) X 2 (Read 3x vs. Retrieved) ANOVA on intrusion responses showed only a marginally significant interaction when the data was not conditionalized on errorless retrieval in Phase 1, $F(1,38) = 3.57$, $p < .07$, $\eta_p^2 = .09$.

Table 1

Demographic data in Experiments 1 and 2.

Experiment and Group	N	Age M(SD)	Vocab. M(SD)	Health[a] Mdn(Range)	Education[b] Mdn(Range)
Experiment 1					
Young Adults	18	19.8(1.5)	33.5(2.0)	4(3-5)	3.0(2-4)
Older Adults	18	76.0(6.2)***	34.9(2.2)	4(2-5)	3.0(1-7)*
Experiment 2					
Young Adults	20	19.5(1.1)	32.8(2.9)	4(3-5)	3.0(3-3)
Unmatched Older Adults	20	75.9(5.7)***	35.6(3.1)**	4(2-5)	3.5(1-7)
Matched Older Adults	20	76.0(8.0)***	34.6(3.5)	4(1-5)	3.0(2-7)

Notes. Vocab.=Vocabulary Subtest of the Shipley Institute of Living Scale. [a]Self-rated health: 1=Poor, 2=Fair, 3=O.K., 4=Good, 5=Excellent. [b]Education levels: 1=Some High School, 2=High School Diploma or GED, 3=Some College, 4=BA or BS, 5=Some Post-Graduate, 6=Master's Degree, 7=Doctoral Degree. Asterisks represent significant differences compared to young adults in the same experiment, with *t*-tests used for Age and Vocab., and Mann-Whitney U used for Health and Education. *p<.05, **p<.01, ***p<.001.

Table 2

Means and standard deviations of the probability of correct recall in Experiment 1.

	Condition		
Age	Read 1x	Read 3x	Retrieved
		Means	
Young	.72	.83	.97 (.95)
Older	.59	.67	.87 (.82)
		Standard Deviations	
Young	.13	.08	.03 (.04)
Older	.11	.11	.12 (.13)

Note. Parentheses represent probability of correct recall without conditionalizing on errorless retrieval in the initial study phase.

Table 3

Illustrative examples of conditions in Experiment 2. On the Final Test, the correct response was the word pair from Phase 2, and the intrusion response was the word pair from Phase 1.

	Condition				
	Read 1x	Read 3x	Retrieved	Guessing	Baseline
Phase 1	sweet-tooth	judge-jury	knee-bend	rabbit-hole	
		judge-jury	knee-b_n_	rabbit-h__e	
		judge-jury	knee-b_n_	rabbit-h__e	
Phase 2	sweet-taste	judge-just	knee-bone		anchor-stop
Final Test	sweet-t__t_	judge-ju__	knee-b_n_	rabbit-h__e	anchor-s__p

Note. Note that particular stimuli varied across participants because of counterbalancing.

Table 4

Means and standard deviations of the probability of correct recall and intrusions in Experiment 2.

	Condition						
Age	Read 1x	Read 3x	Retrieved		Guessing		Baseline
			Mean Correct Recall				
Young	.50	.47	.52	(.53)	.14	(.15)	.67
Older Unmatched	.46	.37*	.28***	(.30)***	.08	(.13)	.58*
Older Matched	.51	.42	.33**	(.36)**	.10	(.14)	.67
			Mean Intrusions				
Young	.42	.47	.45	(.44)	.80	(.79)	.19
Older Unmatched	.44	.53	.69**	(.64)**	.87	(.77)	.22
Older Matched	.41	.50	.61*	(.57)*	.85	(.78)	.19
			Standard Deviation of Correct Recall				
Young	.16	.14	.24	(.23)	.13	(.14)	.14
Older Unmatched	.11	.11	.16	(.15)	.06	(.09)	.12
Older Matched	.12	.16	.17	(.15)	.06	(.09)	.11
			Standard Deviation of Intrusions				
Young	.14	.16	.25	(.24)	.18	(.18)	.11
Older Unmatched	.10	.12	.17	(.17)	.09	(.14)	.08
Older Matched	.11	.15	.16	(.17)	.09	(.12)	.07

Note. Parentheses represent probabilities without conditionalizing on errorless retrieval in the initial study phase. Symbols indicate differences compared to the corresponding mean from younger adults: *p<.05, **p<.01, ***p<.001.