Psychology and Aging

The Role of Forgetting Rate in Producing a Benefit of Expanded Over Equal Spaced Retrieval in Young and Older Adults

Geoffrey B. Maddox, David A. Balota, Jennifer H. Coane, and Janet M. Duchek

Online First Publication, April 4, 2011. doi: 10.1037/a0022942

Psychology and Aging

0882-7974/11/$12.00 DOI: 10.1037/a0022942

The Role of Forgetting Rate in Producing a Benefit of Expanded Over Equal Spaced Retrieval in Young and Older Adults

Geoffrey B. Maddox, David A. Balota, Jennifer H. Coane, and Janet M. Duchek
Washington University in St. Louis

The current study examined the effects of two manipulations on equal and expanded spaced retrieval schedules in young and older adults. First, we examined the role that the type of expansion (systematic vs. nonsystematic) has in producing a benefit of expanded retrieval. Second, we examined the influence of an immediate retrieval attempt to minimize forgetting after the original encoding event. It was predicted that including multiple retrieval attempts with minimal intervening spacing (best accomplished in a nonsystematic retrieval schedule) would be necessary to produce a benefit of expanded retrieval over equal spaced retrieval for older adults but not young adults due to age differences in working memory capacity. Results from two experiments revealed that the presence of an expanded over equal spaced retrieval benefit is modulated by the extent to which the spacing conditions minimize forgetting in the early retrieval attempts in the spaced conditions. As predicted, these conditions differ substantially across young and older adults. In particular, in older adults two intervening items between early retrieval attempts produce dramatic rates of forgetting compared to one intervening item, whereas younger adults can maintain performance up to five intervening events in comparable conditions. Discussion focuses on age differences in short term forgetting, working memory capacity, and the relation between forgetting rates and spaced retrieval schedules.

Keywords: memory, aging, expanded retrieval, spacing effect

There is substantial evidence that episodic memory performance declines in later adulthood (Balota, Dolan, & Duchek, 2000), and so there has been considerable research on the effectiveness of mnemonic techniques that can be used by older adults to minimize this loss. One such technique, spaced retrieval, has been shown to be effective with healthy older adults (see Balota, Duchek, & Logan, 2007, for a review), in addition to Alzheimer's patients (e.g., Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Camp, Foss, Stevens, & O'Hanlon, 1996) and amnesiacs (Schacter, Rich, & Stamp, 1985). There are multiple spaced retrieval schedules that one could use, including equal spacing of retrievals (e.g., each retrieval event after three intervening items) or expanded retrieval (e.g., a retrieval event after one, three, and then five intervening items). There are two questions addressed in the present paper. First, are young and older adults differentially sensitive to various spacing schedules? Second, what are the constraints that should be implemented to maximize the benefits of spacing? We first turn to a brief review of the literature indicating why one would expect young and older adults to be differentially sensitive to various spacing schedules, before reviewing the constraints about schedules of spacing in the extant spaced retrieval literature.

One factor that may modulate the benefits of spaced retrieval between young and older adults is the amount of spacing that occurs between the initial study trial and early retrieval attempts. According to the study-phase retrieval explanation of the spacing effect, the spacing effect is greater for items that are recognized as repetitions on their second presentation (e.g., Braun & Rubin, 1998; Greene, 1989; Thios & D'Agostino, 1976). Thus, the optimal spacing interval may reflect a delicate balance between maximizing spacing while ensuring successful retrieval. In addition, the benefit of increasing interval size may reflect Bjork's (1994) concept of *desired difficulties*. Specifically, Bjork suggested that the memory benefit conferred by retrieval will be greatest when the item is difficult but not impossible to remember. In this light, the degree to which an item persists across spacing intervals may differ between young and older adults due to age differences in both episodic and working memory (e.g., McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010; Park et al., 1996; Salthouse, 1991).

Interestingly, there are surface similarities between spaced retrieval and complex span tasks. For example, many complex span tasks used to assess working memory (e.g., operation span, Turner & Engle, 1989) involve spacing stimuli that are to be remembered for a later test. For example, in the operation span task, participants must learn the serial order of unrelated words that are each separated by a math verification problem. One strategy for learning the to-be-remembered words is to retrieve the previously learned

Geoffrey B. Maddox, David A. Balota, Jennifer H. Coane, and Janet M. Duchek, Department of Psychology, Washington University in St. Louis.

Jennifer H. Coane is now with the Department of Psychology at Colby College.

This work was supported by NIA Training Grant AG00030 and PO1 AGO3991. Portions of this research were previously presented at the Cognitive Aging Conference, 2010, in Atlanta, Georgia. Our thanks to Sarah Moynan Chan, Mitra Haeri, Ricki Diamond, and Anchana Dominic for assistance with various aspects of this project.

Correspondence concerning this article should be addressed to Geoffrey B. Maddox, Department of Psychology, Box 1125, Washington University in St. Louis, St. Louis, Missouri 63130. E-mail: gmaddox@wustl.edu

words when each new word is presented, which is consistent with the study-phase retrieval mechanism proposed to underlie the spacing effect. However, once working memory capacity is reached, it will be difficult to retrieve all of the previously presented items successfully. Similarly, one can imagine a typical spaced retrieval task in which retrieval events occur beyond an individual's working memory capacity. As a result, the stimuli are less likely to be successfully retrieved. Hence, it may be beneficial to include minimal spacing between early retrieval attempts such that retrieval occurs within working memory capacity. This will help ensure the item is well learned so that it can be retrieved from long-term memory rather than working memory after longer spacing intervals.

There is substantial evidence which suggests that age-related differences in working memory capacity underlie age-related differences in episodic memory (e.g., Mayr & Kleigl, 1993; McCabe et al., 2010; Park et al., 1996; Salthouse, 1994; Salthouse & Babcock, 1991). In fact, recent evidence suggests that age-related differences in working memory may account entirely for declines in episodic memory. McCabe et al. (2010) assessed young and older adult working memory capacity, executive functioning, processing speed, and episodic memory using factor analysis and structural equation modeling and reported a powerful role of working memory independent of processing speed. Specifically, results revealed a significant decline in working memory capacity with age that significantly correlated with the observed age-related decline in episodic memory. When differences in working memory capacity were controlled, the age effect in episodic memory performance was eliminated.

Thus, if one considers (a) the benefits of spaced retrieval, (b) the evidence indicating that optimal spacing demands successful reminding of an earlier presentation, and (c) that there is a substantial literature on age-related differences in working memory that have been recently linked to age-related changes in episodic memory, one might expect that older adults may benefit from shorter spaced retrieval intervals than younger adults. Of course, our arguments are not specifically dependent upon age-related differences in working memory, per se, but any difference in age-related mechanisms that produce differences in retrieval of an earlier event could be proposed here. Given the extant accumulating literature, we believe that age-related differences in working memory are likely candidates to account for age-related sensitivity to early spacings.

Before turning to the specific predictions, it is necessary to briefly discuss the literature on different types of schedules of spaced retrieval. In particular, one type of spacing, called expanded retrieval, may be especially beneficial for older adults, as long as the early spaced retrieval events are sufficiently short for successful retrieval of the initial event.

Expanded vs. Equal Spaced Retrieval

Landauer and Bjork (1978) examined multiple spacing schedules in a seminal study with young adults that tested memory for 12 first and last names assigned to one of four spacing conditions. In the *massed* condition, the study and test trials presented during the acquisition phase occurred consecutively (0–0–0, with zero items occurring between each retrieval attempt). In the *equal* condition, the study and test trials were separated by an equal number of intervening trials (e.g., 5–5–5). In the *expanded* condition the study and test trials were separated by an increasing number of intervening items (e.g., 1–4–10). Finally, in the *contracting* condition the study and test trials were separated by a decreasing number of intervening items (e.g., 10–4–1). Results revealed a significant benefit of expanded retrieval over equal spaced retrieval at the end of the acquisition phase (estimated 10%) and at final recall (estimated 5% difference from their figure). Landauer and Bjork's (1978) second experiment involved minor modifications and produced results similar to those discussed above. Based on these results, Landauer and Bjork argued that gradually increasing spacing intervals enables individuals to maintain high accuracy across longer delays, and thus, expanded retrieval is better than simply spacing after equal intervals.

Recent research, however, suggests that there may be some constraints on the benefits of expanded over equal spaced retrieval (when average spacing is controlled), and in some cases, there may be no significant difference between conditions (e.g., Cull, 2000) or even a benefit of equal spaced over expanded retrieval (e.g. Carpenter & DeLosh, 2005; Karpicke & Roediger, 2007). Importantly, these mixed results have been obtained with both young and older adults (i.e., Balota et al., 2006; Logan & Balota, 2008).

A review of the literature (available from the authors) reveals that there may be three factors that significantly contribute to the presence (or absence) of the expanded retrieval benefit. First, the number of retrieval attempts during acquisition appears to be important in producing benefits of expanded over equal spaced retrieval. Of the experiments that failed to produce an expanded retrieval benefit (defined here as better performance on a final test relative to equal spaced retrieval), a vast majority have only included three retrieval events (79%, 26 of 33 experiments)[1]. Additional retrieval attempts may be needed to successfully shape the trace to be retrieved after longer delays in the expanded retrieval condition. Second, it appears that the average spacing interval between repeated retrieval attempts is also important. That is, most studies have used relatively short spacing intervals, and this was particularly true of the 33 studies that have failed to find a benefit of expanded over equal spaced retrieval (M = 3.66 spacing units) compared with nine studies that did obtain this benefit (M = 4.78 spacing units). Finally, the type of expansion appears to modulate the presence or absence of an expanded retrieval benefit. Here we distinguish between systematic and nonsystematic expansion. *Systematic expansion* refers to expanded retrieval intervals that increase by a constant amount (e.g., 2–4–6–8, a constant increase of 2), whereas *nonsystematic expansion* occurs when the interval increases variably (e.g., 1–2–6–8, a nonsystematic increase of 1, 4, and 2) and is often characterized by initial short retrieval events to ensure the trace was well encoded. The use of nonsystematic expansion more frequently produces a benefit of expanded retrieval (44%, four of nine experiments) than systematic expansion (15%, five of 33 experiments). Thus, an

[1] Importantly, Balota et al. (2006) included five retrieval attempts, but two of these attempts occurred immediately after encoding and may not have conferred a true spacing benefit. Additionally, Karpicke and Roediger (2007) included a fourth retrieval attempt in their third experiment, but the inclusion of this fourth attempt produced a schedule combining expanding and contracting intervals.

important contributor to finding an expanded over equal spaced retrieval benefit might be the use of a nonsystematic schedule to establish a stronger memory trace early in the retrieval schedule that then benefits from greater expansion at later retrieval attempts.

Current Study

Based on the review, it is clear that further research is necessary to explore the variables critical in producing a significant benefit of expanded retrieval, especially in light of the practical implications for young and older adults. Past research regarding age differences in working memory and our review of the spaced retrieval literature suggest that the structure of early spacing intervals may modulate memory performance across retrieval attempts during acquisition and at final recall, and importantly, the optimal spacing structure may differ between young and older adults.

Specifically, optimal spacing involves a balance between successful retrieval and maximal delays (e.g., Greene, 1989; Thios & D'Agostino, 1976). Because of their reduced ability to maintain information across intervening items, as reflected by the literature on age-related changes in working memory capacity (e.g., McCabe et al., 2010; Park et al., 1996; Salthouse, 1991), older adults may need a shorter spacing interval between early retrieval attempts for initial retrieval to be successful, compared to younger adults. We believe that early successful retrieval is especially important because as the trace becomes well encoded in long term memory, retrieval success is less dependent on working memory capacity, and as a result, spacing intervals may then expand beyond this capacity limit. This prediction also naturally follows from our review of the importance of nonsystematic retrieval in producing the expanded retrieval benefit. That is, it is critical to have the early retrieval attempts be minimally spaced (most often produced in nonsystematic expansion) to induce significant trace strengthening immediately after the initial encoding event, which then allows for significant spacing between later retrieval attempts. To address this directly, Experiment 1 includes a systematic retrieval schedule, whereas Experiment 2 includes a nonsystematic retrieval schedule to address whether nonsystematic expansion is more likely to lead to an expanded retrieval benefit.

Additionally, our review of the literature suggested two other factors that may increase the probability of obtaining a benefit of expanded over equal spaced retrieval. These are the number of retrieval attempts and the average spacing across retrieval attempts. Thus, in the present experiments we included five retrieval attempts in both the expanded and equal spaced conditions and a relatively larger average spacing of four (Experiment 1) or five (Experiment 2) units to maximize the opportunity of obtaining an expanded retrieval effect.

Experiment 1

As discussed above, one way to increase the benefit of expanded over equal spaced retrieval may be to increase the number of retrieval attempts and the amount of spacing. In Experiment 1, a greater number of retrieval attempts (five attempts) was included than in all past studies, with the exception of Balota et al. (2006). Importantly, we also included an immediate (i.e., massed) retrieval attempt to ensure that both young and older adults initially encoded the items and that this retrieval event occurred within working memory capacity of both age groups. Hence, this experiment included massed (0–0–0–0–0), systematic expanded (0–2–4–6–8), and equal (0–5–5–5–5) conditions for both young and older adults.

Method

Participants. Young adults (N = 30; mean age = 19.75, SD = 1.07, range = 18–22; mean education = 13.92 years, SD = 1.21) were undergraduates at Washington University in St. Louis and received partial course credit or monetary remuneration for their participation. Older adults (N = 30; mean age = 79.2, SD = 7.69, range = 63–91; mean education = 15.14 years, SD = 2.61) were healthy, community dwelling adults and received monetary compensation for their participation.

Materials. A set of 180 weakly associated word pairs was developed, of which 35 had been used previously (Balota et al., 2006). The remaining pairs matched the original pairs in length and frequency. These pairs are considered weak associates because, although the forward and backward associative strength is low ($M \leq .01$; Nelson, McEvoy, & Schreiber, 1998), they often share some features (e.g., *whiskey-water*) or can be easily combined to create a meaningful image or sentence (e.g., *horse-jumped*). The experimental pairs were divided into sets of eight, and each set was counterbalanced across the critical conditions.

Design and procedure. Table 1 displays an example of the list structure. Importantly, the average number of intervening items between retrieval attempts was matched between the equal and expanded retrieval conditions, as was the total number of intervening items between the study event and the last retrieval event. The total number of trials in the encoding phase was 263, of which 24 were encoding trials (eight in each of the experimental conditions), 120 were retrieval practice trials (five for each of the 24 items), eight were primacy and recency trials, and 111 were filler trials. Filler trials were nontested word pair associates meeting the

Table 1
Partial Schedules for Three Different Spacing Conditions

Massed	Equal interval	Expanded
—	horse—JUMPED	horse—JUMPED
—	horse — ?????	horse — ?????
—	—	—
—	—	—
—	—	horse—?????
horse—JUMPED	—	—
horse—?????	—	—
horse—?????	horse—?????	—
horse—?????	—	—
horse—?????	—	horse—?????
horse—?????	—	—
—	—	—
—	—	—
—	horse—?????	—
—	—	—
—	—	—
—	—	horse—?????
. . .	. . .	. . .

Note. Examples of three different Spacing Schedules.

same criteria as those used in the critical conditions. These trials were necessary to ensure that spacing was equated between the appropriate conditions while keeping the list as short as possible and included both study and test trials to minimize the number of to-be-remembered word pairs. Participants were tested individually and completed a short practice list prior to starting the two experimental phases.

Acquisition phase. All study and retrieval events were presented on a computer screen. Each study event included the first presentation of a word pair (e.g., *horse-JUMPED*) for 4.5 seconds. Participants were instructed to study the pairs for a memory test. Subsequent test events consisted of a cued recall trial (e.g., *horse-???*) in which the first word of the pair served as the cue and the second word was the target to be recalled. Participants were given 10 seconds to respond aloud, and the experimenter coded their responses. No feedback was given. The acquisition phase lasted approximately 25 minutes.

Final test phase. The final test phase was administered after a 60 minute filled delay in which participants completed other cognitive tasks not reported here. During the final test, the cue word for each critical pair was presented in random order (e.g., *horse-??*). Again, the participant responded aloud, and the response was coded by the experimenter.

Results

Acquisition performance. Mean percent correct during acquisition for the younger and older adults as a function of spacing condition and retrieval attempt is displayed in Figure 1. There are three observations to note in this figure. First, in contrast to the massed condition, performance in the two spaced conditions declined across the first spaced interval (recall that the first retrieval attempt occurred immediately after encoding, and thus is not expected to differ from the massed condition) and then remained relatively stable across subsequent retrieval attempts. This pattern of performance suggests that once information is sufficiently encoded and retrieved, as indicated by successful retrieval after a short delay (i.e., the second retrieval attempt), it remains relatively preserved for the subsequent three retrieval attempts[2]. Young adults show a small, benefit of expanded over equal spaced retrieval across all retrieval attempts. However, older adults show no similar benefit. Second, the decrease in performance from the first to the second retrieval event was relatively similar for both the equal spaced and expanded retrieval conditions. Third, the decline across the first and second retrieval attempts was much more pronounced for older adults (M = .51) than for young adults (M = .22). In light of these points, it appears that the initial forgetting from the first to the second retrieval attempt is the critical difference between young and older adults.

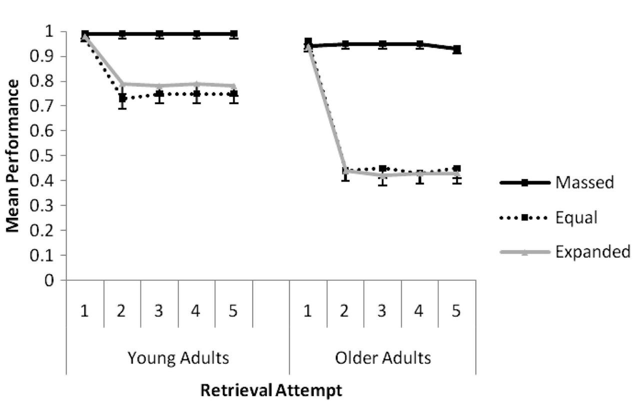

Figure 1. Mean performance during acquisition phase for young and older adults as a function of condition and retrieval attempt. Error bars are 1 standard error below mean performance.

The above observations were supported by a 2 (Age) × 3 (Spacing Condition) × 5 (Retrieval Attempt) mixed factor ANOVA, which yielded main effects of age, $F(1, 58) = 44.10$, $p < .001$, $\eta_p^2 = .43$; spacing condition, $F(2, 116) = 101.35$, $p < .001$, $\eta_p^2 = .64$, and retrieval attempt, $F(4, 232) = 161.43$, $p < .001$, $\eta_p^2 = .74$. Additionally, all two-way interactions were significant (all ps $< .001$). Most importantly, these two-way interactions were qualified by a significant Age x Spacing Condition x Retrieval Attempt interaction, $F(8, 464) = 12.00$, $p < .001$, $\eta_p^2 = .17$ which, as mentioned above and seen in Figure 1, was driven by the different forgetting rates in young and older adults between the first and second retrieval attempts.

Final recall performance. Figure 2 displays final test performance as a function of age and spacing condition. The results are again quite clear. First, there was a significant benefit of both spaced conditions over the massed condition for both young and older adults. Second, there was no difference between equal and expanded retrieval for either age group. The results of a 2 (Age) × 3 (Spacing Condition) mixed factor ANOVA yielded main effects of age, $F(1, 58) = 43.91$, $p < .001$, $\eta_p^2 = .43$, and spacing condition, $F(2, 116) = 42.71$, $p < .001$, $\eta_p^2 = .42$, as well as a significant Age x Spacing Condition interaction, $F(2, 116) = 4.11$, $p < .05$, $\eta_p^2 = .07$. As seen in Figure 2, the interaction is due to a greater benefit in the spaced conditions (averaged across expanded and equal spaced) relative to the massed condition for young adults (M = .29) compared to older adults (M = .16). It is possible that this interaction reflected a floor effect for the older adults in the massed condition. Hence, we conducted a follow-up analysis with the data from a subset of high-performing participants (N = 18) in each age group (i.e., the top three performers out of five in each counterbalancing order for each age group). The results of this analysis revealed significant effects of age, $F(1, 34) = 38.82$, $p < .001$, and spacing condition, $F(2, 68) = 31.66$, $p < .001$. The two-way interaction was no longer significant ($p > .25$) suggesting that the spacing benefit is similar between young and older adults (M = .28 and M = .25, respectively), when one minimizes the influence of floor effects. Finally, in order to more closely focus on the equal and expanded spaced conditions, we excluded the massed condition in a 2 (Age) × 2 (Spacing Condition: Equal,

[2] The conditional probability of correct recall on the fifth retrieval attempt given correct recall on the second attempt was calculated for each age group in the equal and expanded retrieval conditions to further examine the stability and preservation of the trace in both conditions. Results yielded a main effect of age, $F(1, 55)=11.41$, p = .001, such that the conditional probability was higher for young adults (M = .98) than older adults (M = .87). The difference between age groups likely reflects differences in forgetting rates, but importantly, both probabilities are near 1.0 which further supports the conclusion that once initially strengthened, the trace is relatively preserved across subsequent retrieval attempts.

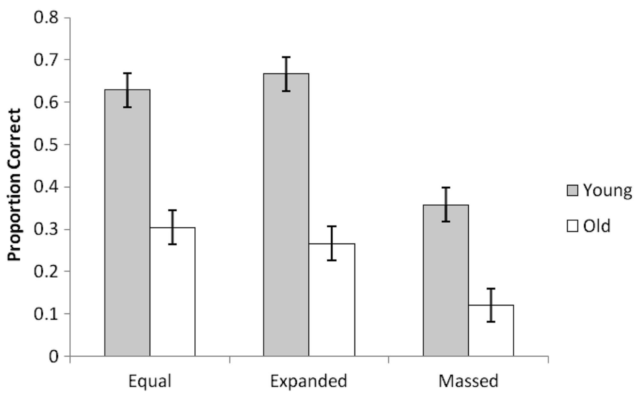

Figure 2. Mean final cued recall performance as a function of age group and spacing condition. Error bars are ± 1 standard error above and below mean performance.

Expanded) mixed factor ANOVA. This analysis yielded a main effect of age, $F(1, 58) = 57.332$, $p < .001$, $\eta_p^2 = .45$. The effect of type of spacing condition and the two-way interaction were not significant which suggests that indeed there were no differences between equal spaced and expanded retrieval.

Discussion

The results of Experiment 1 are clear. First, there was significant forgetting between the first (massed) and second retrieval attempts in both spaced retrieval conditions. Importantly, older adults demonstrated dramatically steeper forgetting than young adults (M = .51 vs. .22, respectively) which is consistent with past research suggesting age differences in forgetting rates at short delays (e.g., Giambra & Arenberg, 1993). This relatively fast forgetting (over only two unrelated intervening items) may reflect an age-related limit on how many items can be retained in a working memory buffer with these materials. Second, there was no significant difference in performance between the expanded and equal spaced retrieval conditions throughout the acquisition phase. Third, there was a significant benefit of spacing over massed practice at final recall, but there was no difference between expanded and equal spacing in either age group. The failure to find a benefit in final recall is consistent with past studies that have used systematic increases in the expanded condition. Of course, this result may be expected because of the equivalent performance between these conditions at the end of the acquisition phase.

Experiment 2

The results of Experiment 1 are consistent with past studies that have used systematic expansion. Hence, Experiment 2 included nonsystematic expansion. A nonsystematic expanded retrieval schedule (0–1–6–8–10) and equal spaced retrieval schedule (5–5–5–5–5) were selected to maximize the likelihood of an expanded retrieval benefit and also to address differences in the forgetting rates of young and older adults. Specifically, in Experiment 1, it appears that the initial retrieval event in both the expanded and equal interval conditions influenced performance on subsequent attempts for young adults by minimizing forgetting across the two-item and five-item spacing intervals, respectively. The same manipulation appears to have yielded relatively little benefit for older adults, who were apparently unable to maintain the item across two intervening items from the first to the second retrieval attempt. To address this possibility, the nonsystematic expansion schedule used in Experiment 2 included a combination of an immediate retrieval attempt and a second retrieval attempt after an even shorter, single item delay. This schedule reflects the minimal amount of spacing that could be used to momentarily shift attention away from the item and then return the target with meaningful spacing and the greatest retrieval success. We predict that older adults will be more likely to be able to maintain the item across the one-item interval, and hence, produce benefits in the expanded schedule.

Method

Participants. Young adults (N = 42; mean age = 19.71, SD = 1.13, range = 18–22; mean education = 13.55 years, SD = 1.90) were undergraduates at Washington University in St. Louis and received partial course credit or monetary remuneration for their participation. Older adults (N = 36; mean age = 77.09, SD = 6.04, range = 65–89; mean education = 15.29 years, SD = 2.37) were healthy, community dwelling adults and received monetary remuneration for their participation.

Materials and procedure. Materials and procedures were identical to those used in Experiment 1 with the exception of the retrieval schedules used in both spacing conditions. The total number of trials in the encoding phase was 224, of which 24 were encoding trials, 120 were retrieval practice trials, eight were primacy and recency trials, and 72 were filler trials. As discussed above, the critical manipulation in Experiment 2 was the use of nonsystematic expansion and the combination of an immediate retrieval attempt followed by a second retrieval event after a short, one-item interval in the expanded retrieval condition (0–1–6–8–10) versus an equal spaced condition with no immediate retrieval attempt (5–5–5–5–5). It is important to note that our nonsystematic expanded retrieval schedule differs from past nonsystematic expanded schedules in one critical way. Although our schedule expands across intervals in absolute duration (0, 1, 6, 8, 10-item intervals), the incremental change across intervals does not always increase as in past studies (i.e. our schedule increases by 1, 5, 2, and 2 items). Critically, expansion occurs in absolute value as well as in incremental change across the first two retrieval attempts, which we believe are critical for producing a significant benefit of expanded retrieval.

Results

Acquisition performance. Mean percent correct during acquisition for the younger and older adults as a function of spacing and retrieval attempt is displayed in Figure 3. There are three observations to note in this figure. First, for both groups, performance in the expanded retrieval condition was higher than performance in the equal spaced condition, and performance in the massed condition was higher than performance in both spacing conditions. Second, performance in the expanded retrieval condition declined from the first to second retrieval attempt and then remained relatively stable across remaining retrieval attempts for both groups. In contrast to Experiment 1, performance in the equal

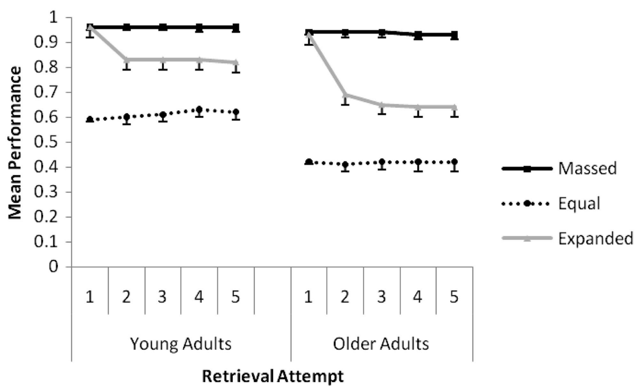

Figure 3. Mean performance during acquisition phase for young and older adults as a function of condition and retrieval attempt. Error bars are 1 standard error below mean performance.

spaced condition did not decline between the first and second retrieval events and in fact remained stable throughout all five retrieval attempts for both groups, although this was to be expected because the first retrieval event in Experiment 1 was functionally massed. Third, older adults showed a much sharper decline in the expanded retrieval condition across the first and second retrieval events than young adults did ($M = .24$ and $M = .13$, respectively). Notably, the overall pattern of performance in the expanded condition again suggests that once information is encoded and successfully retrieved after a short delay it remains relatively preserved for subsequent retrieval attempts[3].

The above observations were supported by a 2 (Age) × 3 (Spacing Condition) × 5 (Retrieval Attempt) mixed factor ANOVA. The results of this analysis revealed significant main effects of age, $F(1, 76) = 15.89$, $p < .001$, $\eta_p^2 = .17$; spacing condition, $F(2, 152) = 183.94$, $p < .001$, $\eta_p^2 = .71$, and retrieval attempt, $F(4, 304) = 40.10$, $p < .001$, $\eta_p^2 = .35$. Planned comparisons revealed significant differences among all three conditions on the fifth retrieval attempt for both age groups (all *p*s < .001). Additionally, all two-way interactions were significant (all *p*s < .001). Most importantly, these two-way interactions were qualified by a significant three-way interaction, $F(8, 608) = 4.64$, $p < .001$, $\eta_p^2 = .06$. As seen in Figure 3, the three-way interaction was driven by relatively stable performance across retrieval attempts in the massed and equal spaced conditions for both age groups, whereas, in the expanded retrieval condition, performance declined more between the first and second retrieval attempt for older adults than younger adults.

Final recall performance. Figure 4 displays the results of the final recall test as a function of age and spacing condition. First, both spacing conditions produced significantly better performance than the massed condition for young and older adults. Second, performance in the expanded retrieval condition was significantly higher than performance in the equal condition. These observations were supported by a 2(Age) × 3(Spacing Condition) mixed factor ANOVA which yielded main effects of age, $F(1, 76) = 30.79$, $p < .001$, $\eta_p^2 = .29$, and spacing condition, $F(2, 152) = 44.52$, $p < .001$, $\eta_p^2 = .37$. Planned comparisons confirmed the significant benefit of expanded over equal spaced retrieval for young adults, $M = .10$, $t(41) = 2.43$, $p < .05$ and older adults, $M = .10$, $t(35) = 2.97$, $p < .005$.

Discussion

The results of Experiment 2 are quite clear. First, and most importantly, the use of nonsystematic expansion produced significantly higher performance than equal spaced retrieval during acquisition and final recall for both younger and older adults. Second, acquisition performance in the expanded retrieval condition again declined between the first and second retrieval events and then remained stable for the remaining retrieval attempts. In contrast with Experiment 1, a similar decline in performance did not occur in the equal spaced condition. Third, this decline in the expanded retrieval condition was steeper for older adults than for younger adults ($M = .24$ and .13, respectively).

Cross Experiment Comparison of Retrieval Schedules

As outlined above, the retrieval schedules used in Experiment 2 were chosen to maximize the probability of an expanded retrieval benefit by using nonsystematic expansion that emphasized the difference in initial encoding between the expanded and equal interval conditions. Indeed for both young and older adults, we were successful in producing an expanded retrieval benefit in both acquisition and final retrieval. Although the average spacing in Experiment 2 was greater than the average spacing in Experiment 1, performance in the expanded retrieval condition of Experiment 2 was better than all of the remaining spaced retrieval conditions in both experiments for both age groups. Thus, a cross-experiment experiment comparison was conducted to provide some insight into the locus of the effect, and also age-related differences in sensitivity to the multiple early retrieval attempts.

Figure 5 displays the performance during acquisition for young and older adults as a function of retrieval attempt, spacing condition, and experiment. In order to explore the influence of schedules across experiments, we conducted a 2 (Age) × 2 (Experiment) × 3 (Spacing Condition) × 5 (Retrieval Attempt) mixed-factor ANOVA. Most importantly, the Age x Experiment x Spacing Condition x Retrieval Attempt interaction was significant, $F(8, 1072) = 5.43$, $p < .001$, $\eta_p^2 = .04$. This four-way interaction was pursued with a series of analyses comparing the effect of spacing condition as a function of retrieval attempt and experiment separately for each age group.

For young adults, the Experiment x Spacing Condition x Retrieval Attempt was significant, $F(4, 280) = 13.06$, $p < .001$, $\eta_p^2 = .16$, which, as shown in the left panel of Figure 5, reflected

[3] Again, the conditional probability of correct recall on the fifth retrieval attempt given correct recall on the second attempt was calculated for each age group in the equal and expanded retrieval conditions. Results yielded main effects of age, $F(1, 73)=6.10$, $p = .02$, and condition, $F(1, 73) = 8.37$, $p = .005$. Similar to Experiment 1, the age difference reflects higher probability for young adults ($M = .95$) than older adults ($M = .89$). The main effect of condition was driven by higher probability in the equal condition ($M = .96$) than in the expanded condition ($M = .88$). This difference reflects the greater degree of spacing early in the equal condition which allowed for greater initial reflecting, and thus, a better indicator of items sufficiently stored in memory compared with a single item interval in the expanded retrieval condition. Importantly, conditional probabilities were again near 1.0 for both conditions, which further substantiates the conclusion that once an item is initially strengthened, the trace remains relatively preserved in memory across subsequent retrieval attempts.

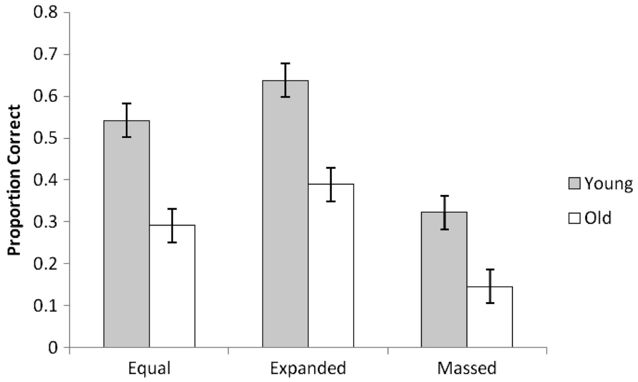

Figure 4. Mean final cued recall performance as a function of age group and spacing condition. Error bars are ± 1 standard error above and below mean performance.

lower performance across all retrieval attempts in the equal spaced condition with no immediate retrieval attempt (5–5–5–5–5) relative to the other three spacing conditions. To further pursue the differences between conditions, performance on the final retrieval attempt of the acquisition phase was compared across conditions. As expected, *t* tests revealed significantly better performance in all three spaced conditions in which an immediate retrieval attempt was included compared to the condition in which spacing occurred after five intervening items (all *p*s $< .01$). Thus, for young adults the initial retrieval (even immediately following the learning trial) is sufficient to efficiently encode the item, which then persists across the multiple retrieval attempts. However, for younger adults, when the first retrieval attempt occurs after a delay of several (i.e., five) items, there is considerable forgetting.

Turning to the older adults, the Experiment x Spacing Condition × Retrieval Attempt was again highly reliable, $F(4, 256) = 13.17$, $p < .001$, $\eta_p^2 = .17$. As shown in the right panel of Figure 5, this three-way interaction reflected superior performance in the expanded condition in which an immediate retrieval attempt was paired with a second retrieval attempt after a short, one-item interval (0–1–6–8–10) relative to other conditions. Importantly, the decline in performance in this condition from the first to second retrieval attempt was not as steep as in the conditions which included an immediate retrieval attempt and a subsequent spacing interval greater than one item (i.e., two and five items). Again, performance on the final retrieval attempt of the acquisition phase was compared across conditions using a series of *t* tests. Performance in the expanded retrieval condition of Experiment 2 (0–1–6–8–10) was significantly better than in the other three spacing conditions (all *p*s $<.02$).

In sum, it is clear that there is one spacing condition during acquisition that is distinct from the three remaining spacing conditions, and importantly, the distinct condition varies as a function of age group. Young adults benefit from an immediate retrieval attempt regardless of subsequent spacing interval (i.e., one, two, and five items) relative to the condition in which spacing occurred immediately. In contrast, older adults produced a very different pattern. Specifically, a single initial retrieval event is not sufficient to boost performance, and indeed performance is quite comparable in the 0–2–4–6–8, 0–5–5–5–5, and 5–5–5–5–5 conditions. For older adults, the critical variable is the presence of multiple retrieval events at a short delay, as in the 0–1–6–8–10 condition. Specifically, the expanded retrieval schedule using nonsystematic expansion (Experiment 2) led to reduced forgetting with the inclusion of an immediate retrieval attempt and a second retrieval attempt after a single item delay. Consistent with our prediction, the inclusion of multiple early retrieval attempts allowed older adults to retrieve items successfully prior to subsequent intervals that extended beyond the limits of working memory capacity.

Mean percent recall at final test is displayed as a function of experiment and spacing condition for young and older adults in Figure 6. Notably, the pattern of performance at the end of the acquisition phase extended into final recall. For young adults, performance in the equal spaced condition of Experiment 2 (5–5–5–5–5) was distinct at acquisition and was reliably worse than performance in both expanded retrieval conditions (*p*s $<.05$) and marginally different than equal spaced retrieval with an immediate retrieval attempt (0–5–5–5–5; $p = .05$, one-tailed) at final recall. For older adults, performance in the expanded condition of Experiment 2 (0–1–6–8–10) was distinct at acquisition and was reliably better than performance in the expanded condition of Experiment 1 (0–2–4–6–8) and the equal spaced condition of Experiment 2 (5–5–5–5–5; *p*s $<.05$) and was marginally different

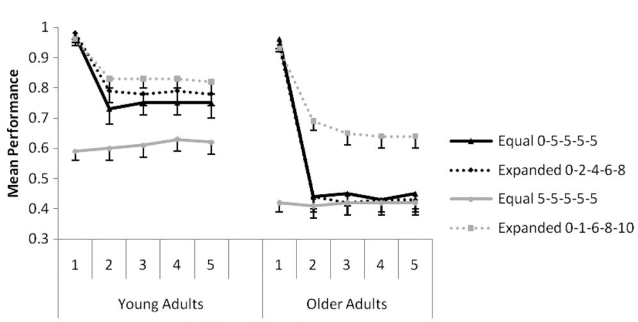

Figure 5. Mean performance during acquisition phase for young and older adults as a function of experiment, spacing condition, and retrieval attempt. Error bars are 1 standard error below mean performance.

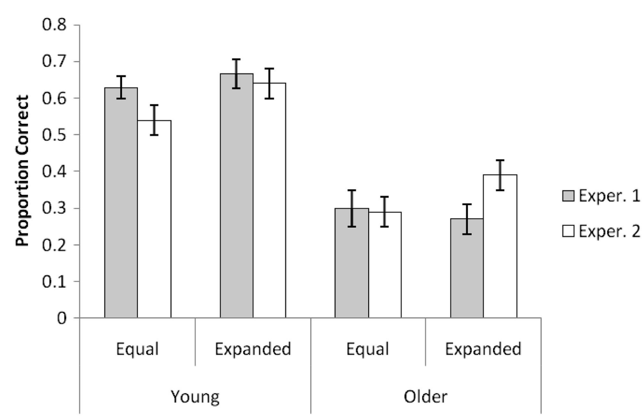

Figure 6. Mean final cued recall performance for young adults as a function of experiment and spacing condition. Error bars are ± 1 standard error above and below mean performance.

than performance in the equal spaced condition of Experiment 1 (0–5–5–5–5; $p = .09$, one-tailed) at final recall. Thus, for young adults, schedules that minimize forgetting by a single immediate retrieval attempt produced a benefit prior to the retention interval that carries over to final test, whereas, older adults benefited from an immediate retrieval attempt only when accompanied by a second retrieval attempt after a short one item interval.

General Discussion

Consistent with our predictions, the spacing of early retrieval attempts modulated the benefit of expanded retrieval. Indeed, our review of the literature indicated that expanded retrieval benefits are more likely to occur when using nonsystematic expansion compared to systematic expansion because there is an increased likelihood of early retrieval events. Experiment 1 examined the benefits of a systematic expanded retrieval schedule (0–2–4–6–8) compared to an equally spaced retrieval schedule (0–5–5–5–5), with five different encoding events. A significant benefit of spaced retrieval was obtained, but there was no difference between expanded and equal spaced conditions for both age groups at final test. Experiment 2 included a nonsystematic increasing retrieval schedule (0–1–6–8–10 vs. 5–5–5–5–5), and produced a reliable benefit of expanded retrieval over equal spaced retrieval at the end of the acquisition phase and at final recall for both younger and older adults.

Importantly, the type of spaced schedule produced very different patterns during acquisition across young and older adults. Specifically, the cross experiment comparisons indicated that young adults produced relatively equivalent performance during acquisition as long as the retrieval schedule included an immediate retrieval event. In contrast, older adults produced a dramatic decrease in performance between the first and second retrieval event in all conditions but one in which the encoding event was followed by retrieval immediately and then again after one intervening item. Remarkably, older adults showed much greater forgetting on average ($M = .42$) than young adults ($M = .19$) in the spacing conditions that introduced more spacing (i.e., two or five items) following the immediate retrieval attempt. We shall now turn to a discussion of possible mechanisms leading to this age-related difference.

First, as discussed in the Introduction, age differences in working memory capacity (e.g., McCabe et al., 2010; Park et al., 1996; Salthouse, 1991) may contribute to the differential benefit of an immediate retrieval attempt followed by a subsequent spacing interval that is relatively short in duration (i.e., one item) for older adults compared with young adults who benefit from an immediate retrieval attempt across greater subsequent spacing intervals (i.e., two or five items). Specifically, reduced working memory capacity may lead to decreases in older adults' ability to maintain an item across the extra intervening item in the 0–2 condition, compared to the 0–1 condition. In contrast, younger adults can not only maintain the information across the 0–2 intervening condition, but also can maintain the items across the 0–5 intervening condition.

Second, the difference between young and older adults may also be related to age-related changes in the benefits of refreshing. *Refreshing* is a component operation of the encoding process within the Multiple-Entry, Modular memory framework that involves maintaining temporary activation of material for easy access and use when cued (Johnson, 1992). In particular, Johnson, Reeder, Raye, and Mitchell (2002) reported that young adults benefited much more from a single immediate retrieval attempt during acquisition than older adults on a final memory test. Likewise in the present study, a single immediate refresh appears to be sufficient to produce relatively high performance in young adults, but in older adults, the refresh needs to also be coupled with a very short spacing of one intervening item.

Taken together, it is likely that immediate retrieval (i.e., refreshing) produces extended activation of an item for young adults but not older adults (Johnson et al., 2002) which then makes the item more accessible for subsequent retrieval across varying delays (i.e., one, two, or five items). Thus, the degree to which the subsequent retrieval event is spaced following immediate retrieval appears to be the critical factor for older adults' memory performance. Consistent with the study-phase retrieval account of the spacing effect (e.g., Braun & Rubin, 1998; Greene, 1989; Thios & D'Agostino, 1976), repetition of an item will only be beneficial if it can be successfully retrieved. We propose that it is likely that both the decreased benefit of refreshing and the decreased working memory capacity in older adults contributed to the pattern of results obtained in the current study. Importantly, for both groups, once the item is well encoded it appears to be maintained at a relatively constant level of performance across the remaining retrieval events. Of course, the different pattern of results between younger and older adults emphasizes the importance of considering the memory abilities of the participants when selecting an optimal spaced retrieval schedule.

In addition to improving memory performance in the current study in older adults, closely spaced retrieval attempts of a given item have been used as a way of improving older adults' recollection (e.g., Bissig & Lustig, 2007; Jennings & Jacoby, 2003). In one recent study (Bissig & Lustig, 2007), participants self-paced their study of words prior to completing a recognition test. Critically, lures were repeated during the recognition tests at varying lags. These lags were relatively short at the beginning of training (e.g., one intervening item) and then increased across training sessions (e.g., 16 intervening items), a manipulation designed to allow participants to learn to distinguish between recollection and familiarity at retrieval. The distinction between recollection and familiarity may occur when participants retrieve the first presentation of a lure item as a way to successfully reject its second presentation rather than rely on simple familiarity. When repetitions of the lure are spaced with few intervening items, retrieval of the first presentation and subsequent rejection of the item as a lure are more likely to occur than when repetitions are spaced farther apart. This process is similar to the study-phase retrieval explanation of the spacing effect; however, in the recollection training paradigm, retrieval of an item's earlier presentation is used to protect against false alarms rather than improve correct recognition (i.e., hits). As a result, there was a small but significant increase in study time across training sessions that was associated with increased performance on the recognition test at greater lags. In this light, it would be intriguing to examine whether recollection training would allow older adults to benefit from longer lags during the first retrieval events in the more standard expanded retrieval paradigm.

The intuitive appeal of expanded retrieval is quite powerful both as a recollection training technique and a general encoding mnemonic. With regard to expanded retrieval's use as a general mnemonic, as noted in the Introduction, it is not entirely clear under what conditions one should expect to find a benefit. We have explored a number of conditions in the present study which indicated at least two important constraints in obtaining an expanded retrieval benefit. First, one should not be compelled to use systematic expansion (i.e., 1–3–5 or 2–4–6–8), which as noted earlier has been typically used in the majority of past studies that have failed to observe a benefit of expanded retrieval. Second, a critical variable in designing spaced retrieval schedules is to include initial encoding that produces a trace that is strong enough to persist across subsequent retrieval attempts. Indeed, the present results clearly indicate that performance at the end of the second retrieval attempt strongly predicts performance at the end of the fifth retrieval attempt, and indeed the influence of spacing on final recall. Third, Storm, Bjork, and Storm (2010) have recently provided evidence that interference between retrieval attempts also appears to be important for producing an expanded retrieval benefit. This may increase the benefit of desired difficulties noted in the Introduction. Finally, and most importantly, one needs to be sensitive to the memory abilities of the particular participants and the targeted population, which will also be influenced by the difficulty of the materials.

Clearly, there is much more work to be done regarding the constraints on which spacing schedules produce the maximum benefits on long-term retention. Based on the relatively stable performance across the third through fifth retrieval attempts in the present study, it will be important to determine what added benefit there is for repeated testing (see Footnotes 2 and 3) within the standard paradigms with multiple retrieval attempts. Specifically, are there diminishing returns of additional retrieval (without feedback) and spacing? Moreover, how does this change across different retention intervals? It is also worth considering the present results in light of the past inconsistencies in the expanded retrieval literature. For example, it is possible that previous studies that have failed to obtain a difference between equal and expanded spaced retrieval may have included retrieval schedules in which the first retrieval attempt in both equal and expanded conditions occurred within or beyond working memory capacity. Although there is still much to be done, the present results clearly indicate that the particular schedule used needs to maximize initial encoding of the stimulus, and this appears to vary quite dramatically across younger and older adults. Hence, in developing an effective spaced retrieval schedule one needs to take into consideration the forgetting rate of the targeted population.

References

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

Balota, D. A., Duchek, J. M., & Logan, J. M. (2007). Is expanded retrieval practice a superior form of spaced retrieval? A critical review of the extent literature. In J. S. Nairne, (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger III* (pp. 83–105). New York: Psychology Press. doi:10.3758/MC.38.1.116

Balota, D. A., Duchek, J. M., Sergent-Marshall, S. D., & Roediger III, H. L. (2006). Does expanded retrieval produce benefits over equal-interval spacing? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging, 21,* 19–31. doi:10.1037/0882-7974.21.1.19

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

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval and time in working memory: Evidence from once-presented words. *Memory, 6,* 37–65. doi:10.1080/741941599

Camp, C. J., Foss, J. W., Stevens, A. B., & O'Hanlon, A. M. (1996). Improving prospective memory task performance in person with Alzheimer's disease. In M. Bandimonte, G. O. Einstein & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 351–367). Mahwah, NJ: USum Associates.

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology, 19,* 619–636. doi:10.1002/acp.1101

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology, 14,* 215–235. doi:10.1002/(SICI)1099-0720(200005/06)14:3::AID-ACP6403.0.CO;2-1

Giambra, L. M., & Arenberg, D. (1993). Adult age differences in forgetting sentences. *Psychology and Aging, 8,* 451–462. doi:10.1037/0882-7974.8.3.451

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15,* 371–377. doi:10.1037/0278-7393.15.3.371

Jennings, J. M., & Jacoby, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation, 13,* 417–440. doi:10.1080/09602010244000390

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience, 4,* 268–280. doi:10.1162/jocn.1992.4.3.268

Johnson, M. K., Reeder, J. A., Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychological Science, 13,* 64–67. doi:10.1111/1467-9280.00411

Karpicke, J. D., & Roediger III, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval promotes long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 704–719. doi:10.1037/0278-7393.33.4.704

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Logan, J. M., & Balota, D. A. (2008). Expanded vs. equal spaced retrieval practice in healthy young and older adults. *Aging, Cognition, and Neuropsychology, 15,* 257–280. doi:10.1080/13825580701322171

Mayr, U., & Kliegl, R. (1993). Sequential and coordinative complexity: Age-based processing limitations in figural transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1297–1320. doi:10.1037/0278-7393.19.6.1297

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology, 24,* 222–243. doi:10.1037/a0017619

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved from http://www.usf.edu/FreeAssociation/

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J., Frieske, D., Zwahr, M., & Gaines, C. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging, 11,* 621–637. doi:10.1037/0882-7974.11.4.621

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2,* 179–183. doi:10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology, 8,* 535–543. doi:10.1037/0894-4105.8.4.535

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763–776. doi:10.1037/0012-1649.27.5.763

Schacter, D. L., Rich, S. A., & Stamp, M. S. (1985). Remediation of memory disorders: Experimental evaluation of the spaced-retrieval technique. *Journal of Clinical and Experimental Neuropsychology, 7,* 70–96. doi:10.1080/01688638508401243

Storm, B. C., Bjork, R. A., & Storm, J. C. (2010). Optimizing retrieval as a learning event: When and why expanding retrieval practice enhances long-term retention. *Memory and Cognition, 38,* 244–253. doi:10.3758/MC.38.2.244

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior, 15,* 529–536. doi:10.1016/0022-5371(76)90047-5

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154. doi:10.1016/0749-596X(89)90040-5

Received June 14, 2010
Revision received January 13, 2011
Accepted January 17, 2011 ■