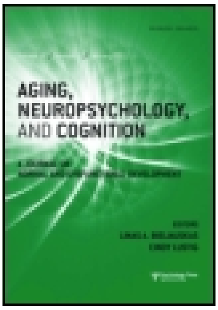


Aging, Neuropsychology, and Cognition
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nanc20

Self control of when and how much to test face-name pairs in a novel spaced retrieval paradigm: An examination of age-related differences

Geoffrey B. Maddox [a] & David A. Balota [a]

[a] Department of Psychology, Washington University, St. Louis, MO, USA

Available online: 25 Jan 2012

Aging, Neuropsychology, and Cognition, 2012, i*First*, 1–24
http://www.psypress.com/anc
ISSN: 1382-5585 print; 1744-4128 online
http://dx.doi.org/10.1080/13825585.2011.640658

Psychology Press
Taylor & Francis Group

Self control of when and how much to test face–name pairs in a novel spaced retrieval paradigm: An examination of age-related differences

Geoffrey B. Maddox and David A. Balota
Department of Psychology, Washington University, St. Louis, MO, USA

ABSTRACT

Although the mnemonic benefit of spaced retrieval is well established, the way in which participants naturally space their own retrieval is relatively unexplored. To examine this question, a novel experimental paradigm was developed in which young and healthy older adults were given control over the frequency and timing of retrieval practice in the context of an ongoing reading task. Results showed that both age groups naturally expanded the intervals of their retrieval practice. When instructed, younger adults but not older adults were better able to employ equal spaced retrieval during retrieval practice. However, even under equal spaced retrieval instructions, young adults included an early retrieval attempt prior to equally spacing their retrieval. Although memory performance was equivalent, secondary task performance was reduced in the experimenter-instructed condition compared with the participant-selected condition. The results overall indicate that both younger and older participants naturally monitor their memory and efficiently use testing to titrate the number and timing of retrieval attempts used during the acquisition phase.

Keywords: Memory; Aging; Testing effect; Retrieval; Metacognition.

This work was supported by NIA Training Grant AG00030, and PO1 AGO3991. Portions of this research were previously presented at Psychonomics, 2010, in St. Louis, Missouri. Our thanks to Mitra Haeri, Anchana Dominic and Jonathan Jackson for their assistance with data collection and to Pooja Agarwal and Jes Logan for their assistance in constructing stimuli. We also wish to thank Mary Pyc for helpful comments on an earlier draft of this manuscript.

Address correspondence to: Geoffrey B. Maddox, Department of Psychology, Box 1125, Washington University, St. Louis, MO, 63130, USA. E-mail: gmaddox@go.wustl.edu

Imagine attending a party and being introduced to a new person, with the hope of learning their name so it can be retrieved at a later point in time. This situation is challenging, because it is essentially a divided attention task that requires learning an arbitrary face–name pair while participating in conversation. Because older adults are relatively poor at learning new names (e.g., Bolla, Lindgren, Bonaccorsy, & Bleecker, 1991; Cohen & Faulkner, 1984, 1986), this is one of their most common memory complaints (Cohen & Faulkner, 1984).

There are a number of ways to approach this social event. For example, one may repeatedly retrieve the name of the person until it feels well-established (*massed retrieval)* and then participate in conversation. One could also switch between retrieval of the name and conversation. In this case, attention could be initially directed toward the conversation with periodic retrieval of the name throughout its duration (*spaced retrieval*). It is this latter strategy that has been promoted in popular media (e.g., Lorenz, 2005) and appears most beneficial for later memory retrieval given past research documenting the benefits of spaced retrieval in both young adults (e.g., Carpenter & DeLosh, 2005; Cull, 2000) and older adults (e.g., Balota, Duchek, & Paullin, 1989; Balota, Duchek, Sergent-Marshall, & Roediger, 2006; Logan & Balota, 2008). Despite this rich literature, no studies to our knowledge have addressed (a) the types of retrieval strategies that individuals choose on their own, (b) the utility of relatively simple instructions by the experimenter to use a strategy, and (c) the benefits of such strategies on later memory performance.

In the present study, we developed a novel paradigm in which participants were given complete control over the frequency and timing of face–name retrieval in the context of an ongoing reading task. In this way, we attempted to bring the demands of learning face–name pairings while engaged in conversation into an experimental setting, and importantly, we aimed to create a task with minimal experimenter influence on strategy selection. To do so, we departed from past metacognitive paradigms in which participants were given partial control of their learning conditions (e.g., choosing to mass, space, or drop an item from future study) but were not free from experimenter-imposed structure (e.g., the spacing interval was pre-determined by the experimenter). Before introducing the experimental paradigm, we will first briefly review the literature addressing these issues.

SPACED RETRIEVAL AND THE BENEFITS OF EQUAL AND EXPANDED SPACING

There is substantial research documenting the benefits of spaced retrieval over massed retrieval in young and older adults (see Balota, Duchek, & Logan, 2007, for a review) with specific emphasis on examining two types of spacing, equal spaced and expanded retrieval. In a seminal paper, Landauer and Bjork

(1978) were the first to report the benefits of expanded retrieval practice relative to equal spaced retrieval practice. In one *expanded retrieval* condition, the number of intervening items between study and retrieval trials increased across events (e.g., the 1–4–10 condition consisted of one item between study and first retrieval, four items between first and second retrieval, and 10 items between second and third retrieval), whereas in the comparable *equal interval* condition, study and retrieval trials occurred with an equal number of items between each event (e.g., the 5–5–5 condition consisted of five items occurring between study and each of the retrieval trials). The results of a 30-minute delayed recall test yielded a small but reliable benefit of expanded over equal interval retrieval.

Expanded retrieval has since become a standard recommendation in memory impaired populations (e.g., Camp, Foss, Stevens, & O'Hanlon, 1996; Schacter, Rich, & Stamp, 1985) and has been discussed in popular media as a way of learning names and faces (e.g., Lorenz, 2005). However, past studies have used experimenter-controlled spacing to improve memory. Thus, it is unclear what spaced retrieval strategies (e.g., massed vs. equal spaced vs. expanded spaced) individuals will adopt on their own. Given the appearance of spaced retrieval strategies in popular media outlets, it may be the case that participants naturally space their retrieval and could implement specific spacing strategies when instructed. However, the evidence from past studies in this literature does not afford a complete answer to this question.

PARTICIPANT BELIEFS REGARDING THE EFFICACY OF SPACING AND TESTING STRATEGIES

There has been growing interest in the metacognitive field regarding the decisions participants make when studying materials. With respect to research that has examined participant choices for massed vs. spaced study, conflicting results have been reported by Benjamin and Bird (2006) and Son (2004). Specifically, Benjamin and Bird reported a spacing preference for difficult items, which is a useful strategy, provided the mnemonic benefits of spaced practice reviewed earlier. In contrast, Son reported the opposite result such that participants preferred to mass difficult items. One critical methodological difference between the Benjamin and Bird (2006) and Son (2004) studies was the presentation rate of items, 5 seconds vs. 1 second, respectively. Toppino, Cohen, Davis, and Moors (2009) pursued this inconsistency by comparing both presentation rates. Spacing preferences for each rate replicated the previous two studies. However, participants reported having trouble perceiving items in the difficult condition at the fast presentation rate which may have led to greater massing of those items. When presentation rate was increased (2.5 vs. 5 seconds) to ensure participants could perceive items in the faster presentation condition (Toppino & Cohen, 2010; Experiment 2), preference for spaced over massed study was greatest for difficult items. Moreover, participants reported on post-experiment questionnaires that they used spaced study trials for the most difficult items.

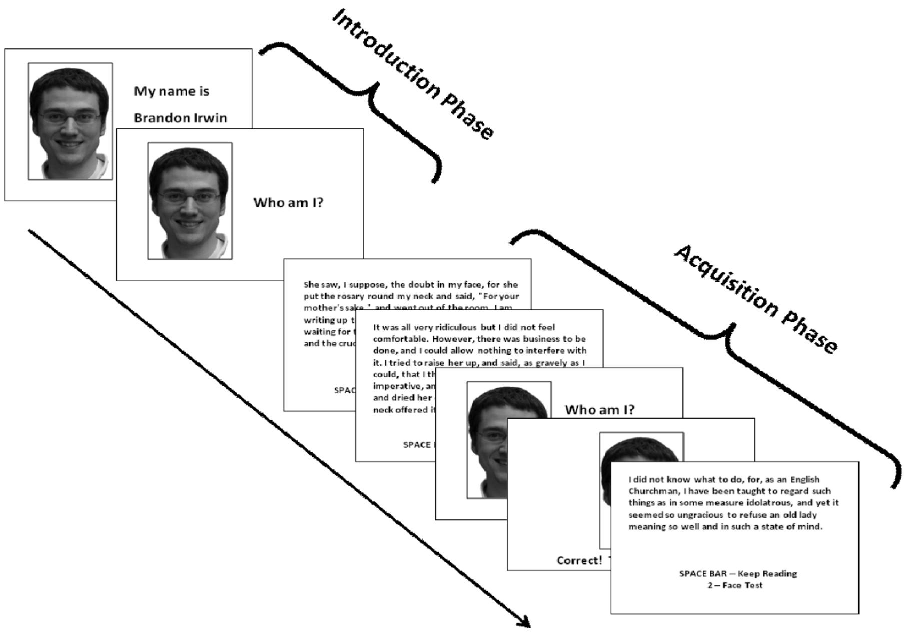
Turning now to situations in which retrieval practice rather than study opportunities are examined, it appears that students are generally unaware of the benefits of testing (e.g., Karpicke & Roediger, 2008). For example, Karpicke, Butler, and Roediger (2009) surveyed student use of quizzing while studying and found that only 8% of those surveyed would quiz themselves to improve memory (also see Kornell & Bjork, 2007). With regard to aging, the ways in which older adults use testing as a mnemonic device in more natural situations has not been directly examined. However, evidence suggests that older adults use self-testing less than younger adults as a learning device in certain tasks (Bottiroli, Dunlosky, Guerini, Cavallini & Hertzog, 2010) and as an assessment of learning prior to taking a test (Murphy, Schmitt, Caruso, & Sanders, 1987). These findings are consistent with other evidence which suggests that older adults spontaneously use many encoding strategies less frequently than young adults (e.g., Dunlosky & Hertzog, 2001; Naveh-Benjamin, Brav, & Levy, 2007). Thus, past research suggests that both age groups may fail to recognize the mnemonic value of retrieval practice as an encoding strategy for long-term retention.

CURRENT STUDY

The present study employs a novel paradigm in which participants learned and maintained face–name associations while managing the simultaneous demands of an ongoing reading task (see Figure 1). Because of the lack of control of natural conversation, we attempted to simulate some of the linguistic demands of a conversation in a face–name learning context by having participants engage in the reading task. Participants controlled both the frequency of retrieval practice and the amount of time between retrieval practice events. This allowed for an assessment of younger and older adult sensitivity to the benefits of testing as a mnemonic device as well as their sensitivity to the benefits of spacing in improving memory. Participants were equally split between two instruction conditions. In the *participant-selected* condition, participants selected their own spacing strategy across three experimental blocks. In the *experimenter-instructed* condition, participants selected their own spacing strategy in the first block and were then instructed on equal spaced and expanded retrieval strategies in the second and third blocks.

There were three aims in the current study. The first aim was to examine how participants use test practice in a situation that has similarities to the demands of learning new face–name pairs in a social context. To the extent that face–name pairs are difficult to remember, participants may prefer spacing their retrieval practice over massing their retrieval practice in a way similar to that reported by Benjamin and Bird (2006) and Toppino et al.

FIGURE 1. Illustration of general method. Following the acquisition phase, participants completed an immediate test phase which consisted of comprehension and face-name tests.



(2009). This pattern of results may be particularly true for older adults given that face–name memory is one of their most common memory complaints (e.g., Cohen & Faulkner, 1984). Alternatively, participants may choose to mass their retrieval practice if there is initial difficulty in remembering the face–name associations and participants perceive they need more time for processing the material. Critically, we extend on past research in the present study, because participants totally controlled when and how frequently they engaged in retrieval practice, within the context of an on-going reading task. This allows participants to dynamically alter their spacing strategies in response to changes in memory strength. In other words, participants may combine massed and spaced retrieval based on how well learned the associations are at a given point in time. Finally, with respect to aging, if timing intervals are dependent on rate of forgetting and participants use this information to modulate testing, then older adults should test earlier than young adults given their clear decline in episodic memory with age (e.g., Balota, Dolan, & Duchek, 2000).

The second aim of the current study was to assess the ability of young and healthy older adults to modify their testing strategies in a more natural, divided attention situation following brief experimenter instructions (i.e., expanded and equal spaced retrieval). Given that popular media suggests spaced retrieval as a technique for learning face–name associations (e.g., Lorenz, 2005) and spaced retrieval techniques have been used to improve memory performance in memory impaired patient populations (e.g., Camp et al., 1996; Schacter et al., 1985), it is important to assess the extent to which participants can apply these different spacing techniques when given explicit instructions.

The third aim was to compare age differences in natural and experimenter-instructed spacing strategies. Specifically, younger and older adults may space their retrieval attempts at different rates given age-related differences in the memorability of face–name pairs. Additionally, there may be age-related differences in the benefit of experimenter-instructed spacing strategies over participant-selected retrieval strategies based on previous research that indicates that older adults are less likely than young adults to spontaneously use beneficial mnemonic strategies in many situations (Devolder & Pressley, 1992; see also Dunlosky & Metcalfe, 2009) and in particular are less likely to use self testing without experimenter instructions (Murphy et al., 1987).

METHODS

Participants

Young adults ($N = 72$; mean age = 20.90, $SD = 3.49$, range = 18–42; mean education = 14.33 years, $SD = 1.70$) were undergraduates at Washington University in St. Louis and received partial course credit or monetary remuneration for their participation. Older adults ($N = 72$; mean age = 76.72, $SD = 6.80$, range = 65–95; mean education = 15.25 years, $SD = 2.58$) were healthy, community dwelling adults and received monetary compensation for their participation. The difference in years of education was significant, $F(1, 134) = 6.09$, $p = .015$.

Materials, Design, and Procedure

A set of 18 faces was selected from the Psychological Image Collection at Sterling (PICS) database. Names were selected from a local telephone book and were medium-frequency to avoid names that were too common or unique. Face–name pairs were equally representative of men and women and young, middle-age, and older adults. Associations were organized into three subsets used separately for each block, as discussed below. Because pilot testing indicated large age-related differences in learning face–name pairs (consistent with the extant literature), younger adults studied six face–name pairs during each block (1 male and 1 female for the young, middle-aged, and older adult stimuli), whereas older adults studied four face–name pairs (1 male and 1 female for the young and older adult stimuli). Passages were selected from three novels (Bram Stoker's *Dracula,* Charles Dickens' *A Message from the Sea*, and H. G. Wells' *The Time Machine*) to serve as reading material in the secondary task. These passages were selected to be relatively easy to comprehend with minimal interference (e.g., character names). The combination of each stimulus subset and passage was counterbalanced across participants.

Design and Procedure

Participants were run individually, and each participant completed three experimental blocks in one of two assigned conditions. In both conditions, instructions in Block 1 emphasized the participant controlled nature of the experiment.

Participant-Selected Strategy Instructions

> In this block you are allowed to test yourself up to four times on the face–name associations. You can use as few or as many of these tests as you wish. You can use them at the beginning, at the end, or you can space them across the reading phase. We trust that you know your memory abilities best, so you should adopt the strategy that works best for you.

Blocks 2 and 3 were different across the two groups of participants. Participants in the *participant-selected* strategy condition were told to continue selecting their own strategy in Blocks 2 and 3. Participants in the *experimenter-instructed* strategy condition were told to use all four retrieval attempts and to equally space or expand their retrieval. The order of equal and expanded retrieval instructions within the experimenter-instructed strategies was counterbalanced across individuals. Prior to providing instructions for the first spacing technique, participants were told that:

> Two groups of researchers believe that different spacing techniques produce better memory performance. We do not have a position in this debate but are interested in examining this question in the current study.

Equal Spaced Interval Instructions

> In this block, please use all four tests and space them equally across the 10-minute reading period. This may benefit memory more than other forms of spacing, because retrieval of the items will be more effortful on the first test than if retrieval occurred shortly after learning the names and faces.

Expanded Retrieval Instructions

> In this block, please use all four tests in the 10-minute reading period. Take the first test shortly after you study the names and faces and then gradually increase the amount of time you wait before taking subsequent tests. This may benefit memory more than other forms of spacing, because it allows you to practice retrieving the names and faces across longer and longer delays.

As seen in Figure 1, each block included an *introduction phase*, an *acquisition phase*, and an *immediate test phase*. Following the third block and a 45 minute delay, participants completed a *final test phase* which included interpolated cognitive tasks. All materials were presented on a computer screen.

Introduction Phase

Participants were instructed to remember the first and last name of each association. Each association was presented individually for 4 (young adults) or 8 seconds (older adults). Encoding time was increased for older adults to ensure performance above floor during the acquisition and final test phases. Participants completed a cued recall trial following the presentation of each association to ensure that the association was accurately encoded.

Acquisition Phase

Following the introduction phase, participants read for 10 minutes and were told to control how quickly they read the passage by pressing the spacebar on the keyboard. They were also told that they could test themselves up to four times on the face–name associations during the 10-minute reading period by pressing the "2" key. Participants were not provided with a timer. On each screen, participants were presented with two to three sentences from the passage and had the options to continue reading or to test themselves on the face–name associations. If participants chose to continue reading, they received the next page of text. If participants chose to take a face–name test, all faces from that block were presented individually, and the participant was asked to generate the appropriate name aloud. Each face was presented for 6 seconds and then the participant's response was coded by the experimenter. Corrective feedback was provided. After the final response was coded by the experimenter, the screen of text the participant was previously reading reappeared which allowed the participant to re-read the material or to continue reading. If participants completed all four face–name tests in a given block, they were only allowed to read for the remaining time.

Immediate Test Phase

Following the acquisition phase, participants answered five true/false questions about the reading, and then completed a cued recall test for the face–name associations. Similar to the face–name tests that occurred during the acquisition phase, each face was presented for 6 seconds and participant responses were coded by the experimenter. However, no feedback was provided.

Final Test Phase

After the third block, participants completed cognitive tasks for 45 minutes and then completed a final face–name cued recall test. Again, participants were presented each face for 6 seconds and required to generate the name which was coded by the experimenter.

RESULTS

Participants Selected Strategies

To assess the spacing strategies participants naturally use, Block 1 performance was collapsed across participant-selected and experimenter-instructed conditions since these were identical during the first block. Given our primary interest in examining the ways in which young and older adults naturally space their retrieval practice, our analyses emphasize the number of tests taken during the acquisition phase and the interval durations separating each test. In addition, we examined recall performance on the final test taken during the acquisition phase as well as on the test that immediately followed the acquisition phase. Finally, we examined performance on the reading comprehension test. Full analysis of acquisition phase recall performance is available via the 'Supplementary' tab on the article's online page (http://dx.doi.org/10.1080/13825585.2011.640658).

Tests Taken

We removed participants who only took zero or one test from subsequent analyses (5 young and 7 older adults, <8% of all participants), because they do not allow us to address the critical aims related to the spacing strategies participants naturally select.[1]

Table 1 presents the proportion of participants as a function of age and number of tests taken during the acquisition phase. Analysis revealed a significant effect of age, $t(130) = 2.03$, $p = .045$, such that young adults ($M = 3.19$ tests) took fewer tests than older adults ($M = 3.46$ tests). As seen in the table, older adults were more likely to use all four tests than young adults which suggests that older adults who took more than one test were sensitive to the benefits of testing (χ^2 values > 15, p values $<.005$).

Spacing Intervals

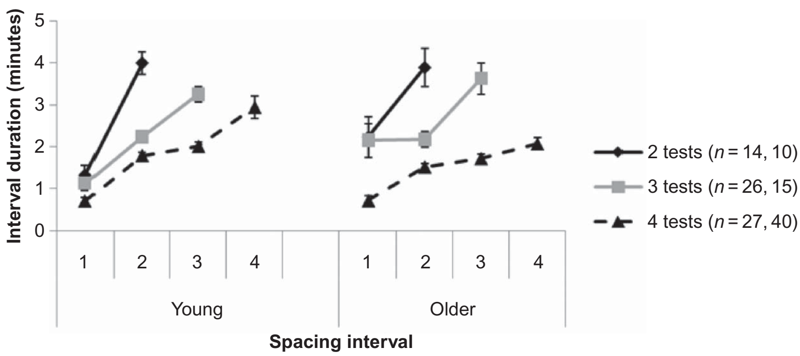
Figure 2 displays mean spacing interval duration as a function of age and tests taken. As shown, both age groups naturally expanded their retrieval practice regardless of number of tests taken during the acquisition phase (i.e.,

[1]When the analysis included all participants, there was a small, nonsignificant effect of age, $t(142) = 0.76$, $p = .451$, such that young adults ($M = 3.01$ tests) took fewer tests than older adults ($M = 3.15$ tests).

TABLE 1. Distribution of participants as a function of age group and number of tests taken during the acquisition phase in Block 1

No. of tests	Young	Older
0	0.03	0.07
1	0.04	0.03
2	0.19	0.14
3	0.36	0.21
4	0.38	0.56
Average	3.01	3.15

FIGURE 2. Spacing interval as a function of age, number of tests taken and spacing interval. Cell sizes included for young and older adults in parentheses, respectively. Error bars are ±1 *SE*.



each spacing interval was longer than the preceding interval). Additionally, young adults expanded their retrieval more than older adults, as reflected by the greater slope for the younger adults than the older adults across tests. Given that participants contributed differentially to interval durations between later retrieval attempts (e.g., between the third and fourth retrieval attempts), the interval durations between study and Test 1 as well as between Tests 1 and 2 were submitted to a 2 (Age) × 2 (Interval) ANOVA. Results revealed a main effect of interval, $F(1, 130) = 86.24$, $p < .001$, $\eta^2_p = .40$, and a significant Age × Interval interaction, $F(1, 130) = 8.20$, $p = .005$, $\eta^2_p = .06$. This interaction reflects the shorter first interval ($p = .13$) and longer second interval ($p = .044$) durations for young adults relative to older adults.

Recall Performance

Table 2 displays recall performance on the final test taken during acquisition and on the immediate test following acquisition as a function of age group and tests taken. One critical question in the current study is how

TABLE 2. Mean (*SE*) recall performance on the final test taken during the acquisition phase and on the immediate test following the acquisition phase during Block 1 as a function of age group and tests taken

		2 tests	3 tests	4 tests
Final acquisition test	Young	0.80 (0.07)	0.85 (0.04)	0.83 (0.05)
	Older	0.63 (0.12)	0.78 (0.07)	0.79 (0.05)
Immediate test	Young	0.93 (0.18)	0.89 (0.14)	0.87 (0.07)
	Older	0.68 (0.11)	0.83 (0.18)	0.76 (0.08)

individuals use testing to achieve a final level of memory performance. Thus, performance on the final acquisition test taken was first submitted to a 2 (Age) × 3 (Tests Taken) ANOVA which failed to yield any significant results (all p values $>.15$). In general, because memory performance was equivalent on the final retrieval attempt regardless of the total number of tests taken, it appears that both younger and older adults are equally efficient at titrating the number of tests taken during learning to achieve similar levels of memory during acquisition.

Turning to performance on the immediate test[2] at the end of acquisition, the data were submitted to a 2 (Age) × 3 (Tests Taken) ANOVA that revealed a significant effect of age, $F(1, 126) = 8.56$, $p = .004$, $\eta^2_p = .06$ such that younger adults ($M = 0.90$) performed better than older adults ($M = 0.76$). Therefore, although we were successful in equating younger and older adults during acquisition, as evidenced by a lack of an age effect on the last test taken during acquisition, there was already a large age effect at the end of acquisition, reflecting an age difference in the forgetting function. It is also interesting to note that the older adults who took four tests performed lower than the older adults who took three tests, which suggests that these individuals may have overall lower memory abilities, and indeed increase the number of tests taken to compensate for this lower memory ability. Additional analyses of acquisition phase tests are available via the 'Supplementary' tab on the article's online page (http://dx.doi.org/10.1080/13825585.2011.640658).

Reading Task Performance

The number of pages read during the reading phase was submitted to a 2 (Age) × 3 (Tests Taken) ANOVA which revealed significant effects of age, $F(1, 126) = 9.10$, $p = .003$, $\eta^2_p = .07$; and tests taken, $F(2, 126) = 10.81$, $p < .001$, $\eta^2_p = .15$. Young adults read more ($M = 29.26$ pages) than older adults ($M = 23.79$ pages), and participants who used two tests

[2]For young adults, final delayed test performance replicated the pattern of results obtained on the end of the block test. For older adults, participants who used four tests performed worse than those who used two and three tests which may reflect the poorer memory abilities of this group.

($M = 32.63$ pages) read more than those who used three ($M = 24.84$ pages) and four ($M = 22.12$ pages) tests.

To accommodate the significant age difference in reading, conditional comprehension scores were calculated to reflect performance for the material that was read during the experiment. Comprehension data were submitted to a 2 (Age) × 3 (Tests Taken) ANOVA. All participants performed relatively well on the reading comprehension test ($M = 0.89$ and $M = 0.88$ for young and older adults, respectively), and there were no significant differences in performance as a function of age or tests taken (all p values $>.15$).

In sum, across age groups all participants naturally expanded their retrieval without any experimenter instructions, and interestingly, younger adults expanded their retrieval at a faster rate than older adults, suggesting that participants were sensitive to their different forgetting functions. In addition, older adults were more likely to take all four tests than younger adults. Performance on the final test taken during acquisition was equivalent across age groups regardless of number of tests taken, which suggests both age groups can monitor their memory and use feedback as a way of titrating the number of tests used during the acquisition phase. Despite equivalent performance across groups on the final test taken during acquisition, young adults performed better than older adults on the immediate test at the end of acquisition, and also on a 45-minute delayed test (see footnote 2). We now turn to the question of whether young and older adults can modify their natural strategies following brief experimenter instructions.

Participant-Selected vs. Experimenter-Instructed Strategies

To address our second aim of comparing performance between the participant-selected and experimenter-instructed conditions, performance in the participant-selected condition was averaged across Blocks 2 and 3 and compared with each of the experimenter-instructed conditions.[3] We make these comparisons separately for (a) equal-instructed retrieval vs. self-selected retrieval and (b) expanded-instructed retrieval vs. self-selected retrieval, because these conditions yielded independent observations. Initial analyses reported for each dependent measure are mixed factor ANOVAs which include all variables, and these were followed by *post-hoc* analyses.

Self-Selected vs. Equal-Instructed Retrieval

Tests Taken

As seen in Table 3, no participants in the experimenter-instructed condition took fewer than three tests (participants were instructed to take all four

[3]Because the presentation order for equal-instructed and expanded-instructed blocks was counterbalanced and collapsed, Blocks 2 and 3 in the self-selected condition were also collapsed and compared separately for each of the instruction conditions.

TABLE 3. Distribution of participants as a function of age group, instruction condition, and number of tests taken during the acquisition phase in Blocks 2 and 3

	Equal		Expanded		Self	
No. of tests	Young	Older	Young	Older	Young	Older
2	0.00	0.00	0.00	0.00	0.20	0.25
3	0.33	0.33	0.31	0.31	0.60	0.25
4	0.67	0.67	0.69	0.69	0.20	0.50
Average	3.67	3.67	3.69	3.69	2.96	3.25

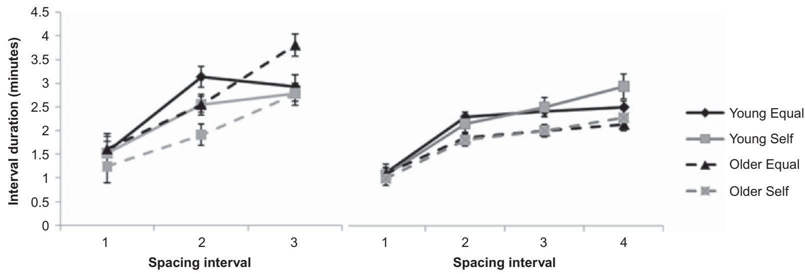
tests); therefore, we analyze only participants who took three or four tests. The mixed factor ANOVA revealed main effects of age, $F(1, 123) = 6.11$, $p = .015$, $\eta^2_p = .05$; and instruction condition, $F(1, 123) = 6.11$, $p = .015$, $\eta^2_p = .05$. Additionally, the Age × Instruction Condition interaction was significant, $F(1, 123) = 6.11$, $p = .015$, $\eta^2_p = .05$, which reflects equivalent testing across age groups in the equal-instructed condition ($M = 3.67$ for both groups) but more testing by older adults ($M = 3.67$) than young adults ($M = 3.25$) in the self-selected condition, $t(53) = 3.35$, $p = .001$.

Spacing Intervals

Mean interval duration is plotted as a function of age, instruction condition, and spacing interval in Figure 3 for participants who used three (left) and four tests (right). First consider data from participants who took three tests. As shown, young adults equally spaced their retrieval across later tests after including an initial, short interval. However, older adults failed to equally space their retrieval, and instead, they expanded their retrieval. The mixed factor ANOVA revealed main effects of interval, $F(2, 100) = 37.93$, $p < .001$, $\eta^2_p = .43$; and instruction condition, $F(1, 50) = 10.67$, $p = 002$, $\eta^2_p = .18$. Additionally, the Age × Interval interaction was significant, $F(2, 100) = 3.97$, $p = .022$, $\eta^2_p = .07$. To further assess this interaction, mean duration for the second and third intervals was submitted to separate *t* tests for young and older adults. Results suggest that young adults were better at following the brief instructions and equally spaced their retrieval across later attempts, $t(11) = 0.78$, $p > .45$, but older adults continued to expand their retrieval, $t(11) = 3.65$, $p = .004$. Although the nonsignificant result for young adults may be due to a lack of statistical power, the difference in absolute duration between intervals was relatively small compared with the 10-minute duration of the reading task ($M = 13$ seconds), especially in comparison to the older adults ($M = 76$ seconds).

Turning to participants who took four tests, young and older adults equally spaced their later tests in the equal-instructed condition following an initial, short interval. Also, there was a slight increase in the fourth interval

FIGURE 3. Spacing interval as a function of age, number of tests taken, spacing interval, and equal-instructed vs. self-selected condition. Error bars are ±1 *SE*.



duration for both age groups in the self-selected condition. The mixed factor ANOVA revealed main effects of age, $F(1, 69) = 11.29$, $p = .001$, $\eta^2_p = .14$; and interval, $F(3, 207) = 76.29$, $p < .001$, $\eta^2_p = .53$; and a marginally significant Instruction Condition × Interval interaction, $F(3, 207) = 2.27$, $p = .081$, $\eta^2_p = .03$, which reflects greater expansion across later tests in the self-selected strategy condition relative to the equal-instructed condition. Although the three-way interaction was not significant, it is noteworthy that separate *t* tests for the second, third and fourth intervals for young and older adults were similar to the pattern obtained for subjects who took three tests, described above. Specifically, there were no differences between interval durations for young adults (p values $>.15$), but there was a significant increase in duration between the second and third intervals, $t(23) = 2.67$, $p = .014$, and between the second and fourth intervals, $t(23) = 2.10$, $p = .047$, for older adults in the equal-instructed condition.

End of Acquisition Test Performance[4]

Mean recall performance is displayed in Table 4 as a function of age group, instruction condition, and tests taken. As shown, young adult performance was relatively equivalent across instruction conditions and number of tests taken. Also, memory performance was relatively equivalent across number of tests taken for older adults in the equal-instructed condition. However, memory performance in the self-selected condition was better for older adults who took three tests than those who took four tests in the self-selected condition and was also higher than performance in the equal-instructed condition.

[4]Similar patterns of performance were obtained on the End of Acquisition Phase and Final Test. Thus, we report data from the End of Acquisition Phase. Young adults remembered more than older adults. Older adults who took four tests in the self-selected condition always performed lower than all other older adult groups.

TABLE 4. Mean (*SE*) recall performance on the immediate test following the acquisition phase during Blocks 2 and 3 as a function of age group, instruction condition and tests taken

		Equal	Expanded	Self
Young	3 tests	0.92 (0.06)	0.94 (0.08)	0.94 (0.05)
	4 tests	0.99 (0.05)	0.92 (0.05)	0.98 (0.08)
Older	3 tests	0.73 (0.06)	0.59 (0.08)	0.92 (0.07)
	4 tests	0.78 (0.04)	0.77 (0.05)	0.56 (0.05)

The mixed factor ANOVA yielded a main effect of age, $F(1, 119) = 25.30$, $p < .001$, $\eta^2_p = .18$. Additionally, the Age × Tests Taken and the Instruction Condition × Tests Taken interactions were significant (p values $< .05$). These two-way interactions were further qualified by a significant three-way interaction, $F(1, 119) = 5.22$, $p = .024$, $\eta^2_p = .04$.

In order to further pursue the three-way interaction, separate 2 (Instruction Condition) × 2 (Tests Taken) ANOVAs were conducted for young and older adults. Analysis of young adult performance revealed no significant main effects or interactions In contrast, analysis of older adult performance revealed a significant effect of tests taken, F(1, 59) = 4.28, p = .04, $\eta^2_p = .07$; and a significant Instruction Condition × Tests Taken interaction, F(1, 59) = 7.66, p = .008, $\eta^2_p = .12$. As seen in Table 4, the interaction reflects near equivalent performance in the equal-instructed condition for participants who used four ($M = .78$) and three tests ($M = .73$), $t(34) = .53$, $p = .597$, but a significant benefit ($t(25) = 3.20$, $p = .004$) in the self-selected condition for participants who used three tests ($M = .92$) relative to those who used four tests ($M = .56$). Similar to the pattern observed in the participant-selected strategy analysis reported earlier for Block 1, it appears that older adults who used all four tests in the self-selected condition may have overall lower memory abilities. Of course, this again converges on the notion that older participants are indeed sensitive to having lower memory performance and hence take more tests.

Reading Comprehension

Results from the mixed factor ANOVA revealed a main effect of instruction condition, $F(1, 119) = 7.09$, $p = .009$, $\eta^2_p = .06$, such that comprehension was better in the self-selected strategy condition ($M = .90$ and $M = .92$, for young and older adults, respectively) than the equal spaced condition ($M = .78$ and $M = .82$, for young and older adults, respectively).

Taken together, the results from the experimenter instructed equal spaced condition indicated that young adults equally spaced their retrieval following an early retrieval attempt, but older adults expanded their retrieval across all tests. It appears that both young and older adults are sensitive to the importance of an initial short retention interval to further strengthen their memory trace. In addition, although young adult performance was statistically equivalent across instruction conditions and number of tests taken on the end of acquisition test, older adults who used four tests in the self-selected condition performed worse than older adults in all other conditions, which is consistent with our earlier Block 1 analysis on participant selected strategies. Hence, in general, it appears that older adults were less sensitive to the equal spacing instructions than younger adults. Finally, reading comprehension was significantly better in the participant-selected strategy condition than the equal-instructed condition, which suggests that the additional monitoring component to follow the equal spacing instructions produced a significant cost to secondary task performance.

Self-Selected vs. Expanded-Instructed Retrieval

Tests Taken

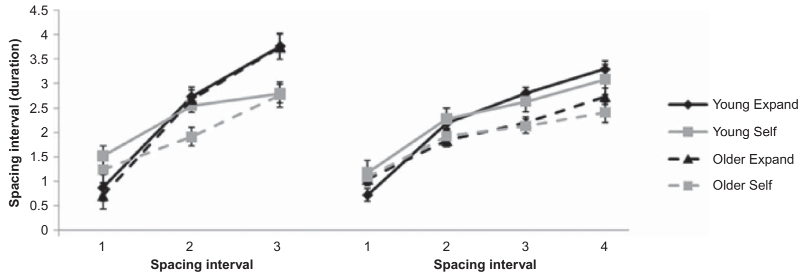
Number of tests taken was submitted to a 2 (Age) × 2 (Instruction Condition) ANOVA which revealed main effects of age, $F(1, 123) = 6.27$, $p = .014$, $\eta^2_p = .05$; and instruction condition, $F(1, 123) = 8.06$, $p = .005$, $\eta^2_p = .06$. Additionally, the Age × Instruction Condition interaction was significant, $F(1, 123) = 6.27$, $p = .014$, $\eta^2_p = .05$, which reflects equivalent testing across age groups in the expanded-instruction condition ($M = 3.69$ for both groups) but more testing by older adults ($M = 3.67$) than young adults ($M = 3.25$) in the self-selected condition (see Table 3).

Spacing Intervals

Figure 4 displays mean interval duration as a function of age, instruction condition, and spacing interval for participants who used three (left) and four tests (right). As shown, both young and older adults who took three tests used nearly identical expanded retrieval schedules when given experimenter instructions. In addition, both age groups expanded their retrieval when selecting their own strategies, but expansion occurred at a slower rate relative to the retrieval schedules used in the expanded-instructed condition. The mixed factor ANOVA yielded main effects of instruction condition, $F(1, 48) = 4.00$, $p = .051$, $\eta^2_p = .08$; and interval, $F(2, 96) = 102.40$, $p < .001$, $\eta^2_p = .68$; and a significant Instruction Condition × Interval interaction, $F(2, 96) = 13.34$, $p < .001$, $\eta^2_p = .22$. As seen in the figure, participants expanded their retrieval naturally but did so more dramatically when given experimenter instructions.

Turning to the participants who took four tests, interval duration increased across all tests for both age groups regardless of instruction condition, but expansion was not as great across later retrieval intervals in

FIGURE 4. Spacing interval as a function of age, number of tests taken, spacing interval, and expanded-instructed vs. self-selected condition. Error bars are ±1 *SE*.



the self-selected condition as it was in the expanded-instructed condition. In addition, older adults included shorter spacing intervals overall. The mixed factor ANOVA yielded main effects of age, $F(1, 71) = 9.27$, $p = .003$, $\eta^2_p = .12$ and interval, $F(3, 213) = 97.22$, $p < .001$, $\eta^2_p = .58$, along with a reliable Age × Interval interaction, $F(3, 213) = 4.29$, $p = .006$, $\eta^2_p = .06$. As seen in Figure 4, young adults expanded more than older adults, and this was particularly true across later spacing intervals (*p* values <.05).

End of Acquisition Test Performance[4]

Mean recall performance at the end of acquisition is displayed Table 4. As shown, there does not appear to be any influence of condition in young adults, whereas for older adults there again appears to be a reliable interaction. The mixed factor ANOVA yielded a main effect of age, $F(1, 119) = 22.65$, $p < .001$, $\eta^2_p = .16$; a significant Instruction Condition × Tests Taken interaction, $F(1, 119) = 6.01$, $p = .016$, $\eta^2_p = .05$; and a reliable three-way interaction, $F(1, 119) = 9.05$, $p = .003$, $\eta^2_p = .07$.

The three-way interaction was further pursued with separate ANOVAs for young and older adults. As seen in Table 4, performance was equivalent across instruction conditions and tests taken for young adults (all *p* values >.50). In contrast, analysis of older adult performance revealed a significant interaction between instruction condition and tests taken, $F(1, 59) = 9.35$, $p = .003$, $\eta^2_p = .14$. This interaction reflected a memory benefit in the self-selected condition for older adults who took three tests over those who took four tests ($M = .92$ and $M = .56$ respectively) but an opposite pattern in the expanded-instruction condition such that older adults who took four tests ($M = .77$) benefited more than those who three tests ($M = .59$), all *p*s < .05. This pattern of results suggests that older adults monitor their memory performance and only use the number of tests they need in the self-selected strategy condition, whereas in the expanded-instructed condition older adults who used all four tests may reflect a group of higher functioning participants as evidenced by their ability to follow directions to take all four tests (i.e., this group may have better monitored the time during the acquisition phase which allowed them to use all four tests before time expired).

Reading Comprehension

The mixed factor ANOVA revealed a marginal effect of instruction condition, $F(1, 118) = 3.59$, $p = .061$, $\eta^2_p = .03$, such that performance in the self-selected condition was better ($M = 0.90$ and $M = 0.92$, for young and older adults, respectively) than in the expanded-instructed condition ($M = 0.86$ and $M = 0.82$, for young and older adults, respectively).

In summary, young and older adults can apply expanded retrieval, which is consistent with the observation that both groups naturally expand their retrieval without any experimenter instructions. Indeed they expanded their study more in the expanded retrieval instruction condition than in the participant selected condition. Similar to the equal-instructed and self-selected comparison, there were no significant differences in memory performance for young adults as a function of instruction condition or number of tests taken at the end of the acquisition phase. However, memory performance for older adults who used three tests in the self condition was better than performance for those participants who used four tests which further supports the conclusion from our earlier Block 1 analysis that the individuals who took more tests appear to have lower memory abilities and took the additional tests to compensate for their lower abilities. Finally, secondary task performance was marginally better in the self-selected strategy condition than the expanded-instructed condition, which again suggests that the addition of time monitoring in the experimenter instructed conditions appears to produce a cost to reading comprehension.

GENERAL DISCUSSION

The current experiments used an approach that diverged from past spaced retrieval and metacognition paradigms to study how young and older adults naturally space their retrieval practice and the degree to which they can successfully apply and benefit from experimenter-instructed spaced retrieval strategies. We shall now turn to the three goals of the present study.

Participant-Selected Spacing Strategies

With respect to the first aim, there were a number of novel observations that afford help in understanding the natural spacing strategies that participants use. In contrast with past research that allowed participants to drop items from further study, or restudy items once either immediately or after a delay (e.g., Benjamin & Bird, 2006; Son, 2004; Toppino et al., 2009), the current study allowed participants to test themselves up to four times in whatever spaced or massed fashion they preferred. In this way, we could examine whether participants optimally space testing events when given *full* control of their spacing intervals. Indeed, both young and older adults expanded their retrieval without experimenter instruction to do so, and this finding was observed for all participants regardless of the number of tests they took during the learning phase.

The participant-selected strategy of expanded retrieval is consistent with current theory regarding the benefits of spaced retrieval, particularly with regard to the concept of *desirable difficulty* (Bjork, 1994). Bjork proposed that the benefit of spaced retrieval is maximized when retrieval of the correct answer is effortful but still successful. With each successful retrieval attempt, the strength of the item increases and the subsequent spacing interval must be increased to obtain desirable difficulty in retrieval. Additionally, the use of an expanded retrieval schedule vs. other forms of spaced retrieval extends past metacognitive research regarding participant preferences for spacing vs. massing material. A preference for massing study events for difficult items has been previously reported by Son (2004). This preference was further examined by Toppino et al. (2009) in a study which concluded that the degree to which material is initially encoded (as a consequence of presentation rate) influences the preference for spacing. Given that the current study used arbitrary face–name pairings and that these associations are one of the most frequent memory complaints of older adults (Cohen & Faulkner, 1984), including an early retrieval attempt may reflect a need for additional processing time shortly after initial encoding. As the associations are better acquired, spacing intervals increase in duration, a finding consistent with past studies in which participants only had two study opportunities (Benjamin & Bird, 2006; Toppino et al., 2009).

It is also important to note that all participants obtained a similar level of performance at the end of the acquisition phase (regardless of the number of retrieval attempts taken), which suggests that participants were able to monitor their memory performance and use feedback as a way of titrating the number of retrieval attempts used in the current dual-task paradigm. This finding extends past metacognitive research which suggests that memory monitoring accuracy is relatively intact in later adulthood (e.g., Connor, Dunlosky, & Hertzog, 1997; also see Hertzog & Dunlosky, 2011). Importantly, there were also some intriguing age differences. For example, the results indicated that older adults were more likely than younger adults to use all four tests when allowed to choose their own strategies. In addition, older adults expanded at a slower rate than younger adults (i.e., shorter increasing intervals across tests), which would be predicted based on their faster forgetting. It appears that young and older adults understand the benefits of repeated testing and spacing and engage in retrieval strategies that are tuned to different forgetting rates to benefit final acquisition.

The results from the current study also help qualify the inferences drawn from past literature suggesting that young adults fail to fully recognize the mnemonic benefits of testing (e.g., Karpicke et al., 2009; Kornell & Bjork, 2007). If we first consider young adults, Karpicke et al. reported that 72% of students who reported using test practice when studying did so as a way of identifying the areas they needed to re-study and only 2% of students did so for the mnemonic benefits of test practice. Indeed, participants in the current study test themselves when given full control, and they may have done so either to assess learning or to benefit from corrective feedback. However, it is important to note we (Maddox & Balota, 2010) have previously reported results from the same paradigm as the one used in the current study, but no corrective feedback was provided after the reading-phase tests. In our earlier study, results revealed that 63% of young adults continued testing after taking the first test even when objective assessment of learning was impossible, i.e., they were not given any feedback. Similarly, 61% of older adults continued to test themselves after taking the first test. Importantly, the average number of tests taken when feedback was not provided was not significantly different between young and older adults ($p = .425$; $M = 2.03$ and $M = 2.28$ tests, respectively). Thus, our earlier results suggest that both age groups appear to be sensitive to the mnemonic benefits of testing even when assessment of learning via feedback is not provided.

Interestingly, there is evidence in the literature suggesting that, compared to younger adults, older adults are less likely to spontaneously engage in efficient encoding strategies (e.g., Dunlosky & Hertzog, 2001; Naveh-Benjamin et al., 2007). It may be that the mere presence of the option to test on the associations during the acquisition phase in the present study was sufficient encouragement for participants to use testing despite explicitly instructing them to use as few or as many tests as they thought necessary for their memory strength. It is important to note, however, that only one older adult (and no younger adults) reported using the cue on the screen as a reminder to use the testing option. Hence, they appeared to be internally monitoring their memory strength for the face name pairs. Furthermore, the present results converge on the observation that some older adults do naturally use testing as a mnemonic device without task cues or explicit experimenter instructions (Bailey, Dunlosky, & Hertzog, 2009). The present study extends this finding by showing that older adults take advantage of the opportunity to test themselves multiple times during learning and separate these tests with increasing amounts of intervening spacing. Importantly, we would argue that subjects appear quite naturally adept at expanding their retrieval in a situation which has similarity to the memory demands in dual task situations, where participants can retrieve and practice to-be-remembered information at their own will. Thus, it appears that participants continue testing when they have not fully learned the material, which further suggests that both young and older adults can naturally monitor their memory and use feedback to titrate the number of retrieval attempts used during the acquisition phase.

Participant Use of Equal Spaced and Expanded Retrieval

Turning now to our second aim, we examined young and older adults' ability to modify their natural spacing strategies following brief experimenter instructions to equally space or expand retrieval. Results from experimenter-instructed strategy blocks suggest that young adults can successfully apply equal spaced retrieval but still include an early, short interval prior to their first test. In contrast, older adults were more likely to expand their retrievals under instructions emphasizing equal spaced retrieval, even beyond a first short retrieval. Thus, it appears that younger adults were more sensitive to the instructional set in modifying their spacing schedules. Of course, because both age groups naturally expand their retrieval (as reflected by the self-selected spaced intervals), it was not surprising that young and older adults successfully applied expanded retrieval (to a larger extent than on their own) when instructed.[5]

Benefits of Participant-Selected vs. Experimenter-Instructed Spacing Strategies

In addition to assessing the ability of young and older adults to apply these techniques, we examined the benefit of these techniques relative to self-selected strategies. First, memory performance was equivalent across experimenter-instructed and self-selected strategy conditions, but secondary task performance was better when participants selected their own strategies relative to applying experimenter instructions. The trade-off in secondary task performance across instruction conditions likely reflects the increased processing demands (e.g., time monitoring) in the experimenter-instructed condition. Of course, this trade-off may be reduced or eliminated if participants are provided increased practice applying these spaced retrieval schedules.

[5]Additionally, a direct comparison of the end of block test performance in the expanded and equal spaced instruction conditions for participants who took four tests (17 young and 19 older adults) revealed a marginal effect of age, $p = .057$, such that performance was higher for young adults ($M = 0.94$) than older adults ($M = 0.80$), and no effect of instruction condition, $F(1, 34) = 1.17$, $p > .29$.

Results from the current study can inform the way in which spaced retrieval techniques are described to young and healthy older adults. Based on the current paradigm, it appears that participants naturally space their retrieval and titrate the number of tests they take during acquisition based on performance monitoring and corrective feedback. Without additional practice applying specific spacing constraints, encouraging participants to use a specific spacing schedule may be detrimental to ongoing task performance (also see Son, 2010). That is, it appears that participants are relatively good judges of when to test, at least within the present paradigm. This should not be surprising given their lifelong experience with their own memory systems. Of course, it is important to extend this work to additional paradigms that may make less salient the mnemonic benefits of testing and spacing.

In summary, the present results have yielded a number of noteworthy effects. First, young and older adults naturally expand their retrieval when provided complete control over the frequency and spacing of such events. Second, it appears that young and older adults monitor their memory and use feedback from retrieval practice to titrate the number of tests taken during acquisition. Third, young adults successfully applied equal spaced and expanded retrieval skills when instructed to do so, but older adults were only successful in applying expanded retrieval skills. Importantly, both young and older adults included an early retrieval attempt that occurred shortly after initial encoding regardless of instruction condition. Fourth, participant-selected and experimenter-instructed strategies produced similar levels of memory performance, but there was a greater cost to secondary task performance when participants were asked to adopt a specific spacing strategy relative to selecting their own strategy.

Although the present results are clear, this study also provides fodder for future studies. First, it would be useful to examine the ability of young and older adults to apply spaced retrieval strategies after receiving multiple training trials with each type of spacing. Second, introducing external support (e.g., a timer or watch) may improve task performance in the memory and reading tasks by reducing the time monitoring component. Third, the current study required participants to test on all face–name associations with each acquisition test. Allowing participants to select test strategies for individual items will provide additional examination and replication of the current results. Finally, although the current paradigm has addressed the ways in which participants choose to test themselves when given greater control over the frequency and timing of such events, introducing more difficult ongoing tasks and manipulating memory load may provide additional ways of examining participant decisions when to test in more natural settings. Importantly, the present novel experimental paradigm affords a framework for pursuing each of these important issues.

Supplementary Material

Supplementary material is available via the ‘Supplementary’ tab on the article’s online page (http://dx.doi.org/10.1080/13825585.2011.640658).

Original manuscript received 2 August 2011
Revised manuscript accepted 8 November 2011
First published online 24 January 2012

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, NY: Psychology Press.

Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag and retention interval. *Psychology and Aging*, *4*, 3–9.

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

Benjamin, A. S., & Bird, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language*, *55*, 126–137. doi: 10.1016/j.jml.2006.02.003

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.

Bolla, K. I., Lindgren, K. N., Bonaccorsy, C., & Bleecker, M. L. (1991). Memory complaints in older adults: Fact or fiction? *Archives of Neurology*, *48*, 61–64.

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

Cohen, G., & Faulkner, D. (1984). Memory in old age: ‘Good in parts’. *New Scientist*, *11*, 49–51.

Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology*, *4*, 187–197.

Connor, L., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*, 50–71.

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, *14*, 215–235.

Devolder, P. A., & Pressley, M. (1992). Causal attributions and strategy use in relation to memory performance differences in younger and older adults. *Applied Cognitive Psychology*, *6*, 629–642.

Dunlosky, J., & Herzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition*, *29*, 247–253.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Los Angeles, CA: Sage.

Hertzog, C., & Dunlosky, J. (2011). Metacogntiion in later adulthood: Spared monitoring can benefit older adults' self-regulation. *Current Directions in Psychological Science*, *20*, 167–173.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General*, *138*, 469–486. doi: 10.1037/a0017341

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*, 966–968. doi: 10.1126/science.1152408

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*, 471–479. doi: 10.1080/09658210802647009

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*, 219–224.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the 'enemy of induction?' *Psychological Science*, *19*, 585–592. doi: 10.1111/j.1467-9280.2008.02127.x

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, UK: 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

Lorenz, K. (2005). Tricks to remembering names. *CNN*. Retrieved from www.cnn.com.

Maddox, G. M., & Balota, D. A. (2010, November). *Equal versus expanded retrieval: An examination of participant- and experimenter-selected spaced retrieval schedules*. Poster presented at Psychonomics, St. Louis.

Murphy, M. D., Schmitt, F. A., Caruso, M. J., & Sanders, R. E. (1987). Metamemory in older adults: The role of monitoring in serial recall. *Psychology and Aging*, *2*, 331–339.

Naveh-Benjamin, M., Brav, T. K., & Levy, O. (2007). The associative memory deficit of older adults: The role of strategy utilization. *Psychology and Aging*, *22*, 202–208. doi: 10.1037/0882-7974.22.1.202

Psychological Image Collection at Sterling (PICS) Database. Retrieved from http://pics.psych.stir.ac.uk/

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.

Son, L. K. (2004). Spacing one's study: evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *30*, 601–604. doi: 10.1037/0278–7393.30.3.601

Son, L. K. (2010). Metacognitive control and the spacing effect. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *36*, 255–262. doi: 10.1037/a0017892

Toppino, T. C., & Cohen, M. S. (2010). Megacognitive control and spaced practice: Clarifying what people do and why. *Journal of Experimental Psychology, Learning, Memory & Cognition*, *36*, 1480–1491. doi: 10.1037/a0020949

Toppino, T. C., Cohen, M. S., Davis, M. L., & Moors, A. C. (2009). Metacognitive control over the distribution of practice: When is spacing preferred? *Journal of Experimental Psychology: Learning, Memory & Cognition*, *35*, 1352–1358. doi: 10.1037/a0016371