AGING, NEUROPSYCHOLOGY, AND COGNITION, 2017
VOL. 24, NO. 6, 703–716
https://doi.org/10.1080/13825585.2016.1251552

The spacing effect in older and younger adults: does context matter?

Katherine E. Bercovitz[a], Matthew C. Bell [a], Patricia M. Simone[a] and Melody Wiseheart [b,c]

[a]Department of Psychology, Santa Clara University, Santa Clara, CA, USA; [b]Department of Psychology, York University, Toronto, ON, Canada; [c]LaMarsh Centre for Child and Youth Research, York University, Toronto, ON, Canada

ABSTRACT
Age-related memory change has been a topic of much investigation in recent years, including spacing benefits and reliance on contextual cues. We manipulated the spacing schedule and the context of learning and observed the effects on long-term recall ability in healthy older and younger adults. After learning Swahili–English word pairs, half practiced immediately (massed) and half practiced 24 h later (spaced) either in the same room or a different room (context) from the initial session. A final recall test 10 days after the practice session occurred in the same room as the first session. Participants in the spaced condition remembered more than those in the massed condition 10 days later. Younger adults remembered more word pairs than older adult participants. Context change eliminated the spacing benefit for both age groups.

ARTICLE HISTORY
Received 1 June 2016
Accepted 15 October 2016

KEYWORDS
Memory; aging; context; spacing effect

Distributed practice (i.e., spaced learning) of verbal information has consistently been shown to improve recall performance in younger adults (for a review, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). This spacing confers a memory benefit when presentations are within a given session (Maddox, Balota, Coane, & Duchek, 2011), separated by many hours (Smith, Glenberg, & Bjork, 1978), and when distributed across many days (Glenberg & Lehmann, 1980). Compared to young adults, older adults have consistently shown difficulties in remembering associative information (e.g., Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). One approach to boost memory performance in older adults is to capitalize on the spacing benefit. Older adults, however, may not benefit from spacing to the same degree as younger adults (e.g., Simone, Bell, & Cepeda, 2013), which may be related to differences in the ability to utilize contextual information in paired-associate learning tasks.

Spacing effect and context

Spacing is when a learning episode includes a delay between two presentations of a to-be-remembered item. The spacing benefit is observed when the item is later recalled

CONTACT Matthew C. Bell mbell@scu.edu

after some delay. Spacing studies differ in encoding session (limited presentations or criteria learning), length of inter-study interval (also known as gaps which can vary between just a few items in item-level studies to hours in list-level studies), retention interval (which can vary between a few minutes to several days), and recall method (free recall, cued recall, recognition). Possibly due to differences in methodology (and thus results), one unified theory of the spacing benefit has not been identified. Many of the theories of the spacing benefit, however, posit an important role of context, with some debate about how it affects learning.

Contextual variability theory explains the spacing benefit in terms of differing contextual cues being available with the passage of time (e.g., Madigan, 1969; Melton, 1970; see Delaney, Verkoeijen, & Spirgel, 2010; for a more contemporary review). According to Estes's (1955) stimulus-fluctuation model (an early model addressing effects of contextual variability), a spacing schedule improves memory performance because additional cues become associated with the target word when there is a gap between target word presentations. Similarly, Melton (1970) proposed that the number of encoded retrieval cues associated with the target increase when items are spaced, not massed, due to changes in context that occur over time. The increase in contextual cues associated with spaced items should then result in enhanced ability to recall the target item at a later time. In other words, the more contextual cues associated with the target, the better the recall. Similarly, the study-phase retrieval hypothesis (Thios & D'Agostino, 1976) also addresses the role of context in spacing studies. It asserts that delayed retrieval following initial encoding is critical to improving long-term memory due to the changing of contextual cues that occur over time allowing additional cues to be associated with the target and thereby assisting recall at a later time. This study-phase retrieval hypothesis emphasizes the importance of successful retrieval following the gap in presentations. If the gap is too long, the target will not be retrieved and the spacing benefit will be lost.

Context changes with the passage of time, as highlighted by spacing benefit theories, but context can also vary with explicit changes of physical features of the environment, and these contextual changes may confer a similar benefit to learning and interact with spacing in different ways. For example, in an item-level spacing study with college students, Verkoeijen, Rikers, and Schmidt (2004) found that contextual similarity (screen background color, image) between spaced presentations of the target resulted in superior performance while contextual change between presentations of target items improved recall of massed items. Verkoeijen et al. concluded that contextual variability theory explains the boost in massed performance and study-phase retrieval hypothesis explains the superior performance of spacing when context is consistent. Similarly, Smith et al. (1978) found that exposing college-aged participants to a novel context (time of day, the building and room, appearance of the experimenter, and sensory modality of stimulus presentation) in the study-phase session improved final test performance in a neutral context. Thus, there is evidence supporting both study-phase retrieval and contextual variability theory of the spacing benefit (Greene, 1989; Raaijmakers, 2003; see also Küpper-Tetzel, 2014; for a review). The important point for our purposes is that both theories emphasize the important role of context.

Spacing effect and age

Contextual cues appear important for the spacing benefit to occur. The role of contextual processing in spacing studies with older adults has not yet been examined in the same detail as with younger, college-aged participants. Older adults can benefit from spacing in both item-level spacing studies (e.g., Balota, Duchek, & Paullin, 1989) and list-level spacing studies (e.g., Simone et al., 2013), but to a lesser degree than younger adults. One possible explanation for this difference in the spacing benefit is that it is the result of differences in utilization of contextual cues by younger and older adults. For example, Light and Singh (1987) suggested that older adults may have a deficit with storing and/or utilizing contextual information. Similarly, Balota et al. (1989) suggested, based on finding performance differences between older and younger adults, that older adults encode fewer contextual cues moment to moment. Balota et al. suggested that the passage of time may not yield the same variability in contextual cues for older as it does younger adults and this difference may negatively impact the spacing benefit in older adults, especially with fewer contextual cues afforded by long spacing gaps and/or changing of the physical context over sessions.

Other work suggests that contextual support is used to facilitate learning and memory in older adults. For example, Craik's (1983, 1986) environmental support hypothesis suggests that because older adults are less able to self-initiate successful encoding strategies, they may actually rely more on environmental cues to assist memory. In other words, contextual stability provides cues to facilitate retrieval for older adults (Anderson, Craik, & Naveh-Benjamin, 1998; Craik & Rose, 2011).

Similarly, Luo, Hendriks, and Craik (2007) found that older adults benefit from contextual enhancement more than younger adults. In their study, when the to-be-learned words were paired with pictures, older adults showed more improvement in recollection compared to younger adults. Thus, while the number of contextual elements available in spaced learning may serve to cue retrieval in both younger and older adults, older adults may actually depend more on those stable contextual cues to remember than younger adults. As a result, performance for older adults should decline as fewer contextual cues are available, including spacing studies if what is happening is that older adults encode fewer contextual cues over the same amount of time as younger adults (Balota et al., 1989). However, Luo et al. (2007) manipulated pairing of words and pictures in an item-level spacing study, and this procedure may affect reliance on context differently than the proposed study.

Current study

The goal of the current study was to examine the effect of substantial context changes on the spacing benefit in younger and older adults and, by doing so, to evaluate the role of context in conferring a spacing benefit. Using a paired-associate memory task (systematically replicating Simone et al., 2013), we manipulated whether cued retrieval practice occurred in the same or different room than the original encoding/testing room and whether participants had no delay or a 24-h lag between initial encoding and the subsequent practice session. Our key-dependent measure was cued-recall performance at a retrieval session occurring 10 days after the practice session.

The study-phase retrieval hypothesis predicts that changing context should have a detrimental effect on the spacing benefit because fewer items will be recalled in the study phase (practice session). Contextual variability theory suggests that context changes during the practice session should enhance learning of the massed items due to association with additional contextual elements compared to a condition where the context does not change.

Contextual stability is, however, important for learning in some situations. Godden and Baddeley (1975), for example, found that participants who encoded and retrieved 36 unrelated words in a list-level study remembered more words when the retrieval was in the same context (above ground or under water) than when retrieval was in a different context. It seems, then, that additional contextual information can be helpful to cue recall, but may in fact also harm learning if the learning becomes too dependent on less important contextual cues.

It is unclear what effect intentional context change should have on the spacing benefit for older adults in a cued recall list-level learning procedure over a meaningful delay in which many contextual cues change simply as a function of time elapsing. Glenberg (1979) suggested that the automatic encoding of contextual information present in multiple memory traces helps cue retrieval. However, if older adults sample less of the environment (Balota et al., 1989), then they are less likely to encode contextual information that is potentially helpful. Even so, some of the context that is encoded may be irrelevant to the stated target task (Campbell, Trelle, & Hasher, 2014). If older adults depend upon those external cues to initiate successful encoding strategies (Craik, 1983, 1986), then a significant context change (i.e., our room change) at the time of reminding (study phase retrieval) should negatively affect performance for older adults in massed and more so in spaced conditions (which has the additional contextual change that occurs over time).

Method

Participants

Older adult participants, recruited from the community in Santa Clara, California ($N = 94$), were healthy, active, and living independently. They were randomly assigned to one of four groups of a 2 × 2 matrix with spacing (massed or spaced) and context (same room or changed room) as the factors. Older adult subjects who were more than 1.5 standard deviations above the mean number of trials required to successfully complete the encoding session (for older adults) were excluded from analysis. Additionally, one older adult's data were excluded from the analysis because of a computer malfunction during the practice session. Thus, data from 86 older adult participants were analyzed. Demographic data for some of the older adults were lost; however, the sample of participants were relatively similar in age and education. For the remaining participants, the mean age was 68 years ($SD = 6$; $N = 73$), and mean education level was 18 years ($SD = 2$; $N = 67$). For each age group, there were no significant differences between the experimental groups for either age or education level.

Younger adults were recruited from a pool of undergraduate students at Santa Clara University ($N = 95$) for which the mean age is approximately 19 years, and most were either first or second year students. One participant's data were excluded because of a computer malfunction that prematurely ended the encoding session. Participants were predominately female (69% older adults, 80% younger adults). Older participants were paid and younger participants earned course credit for their participation. Sample size was selected based on our estimate of a large effect size, given previous related research and meta-analyses (Bell, Kawadri, Simone, & Wiseheart, 2014; Cepeda et al., 2009, 2006; Donovan & Radosevich, 1999; Simone et al., 2013). We estimated that 100 participants total would provide 95% power to detect our predicted large effect size, and our actual sample size provided 80% power to detect a medium effect size.

Materials and procedure

Four separate rooms were used, all of which were approximately the same size. Two rooms were located on the first floor and were not decorated. The two rooms located on the second floor were noticeably different from the first-floor rooms and were made more distinctive with a fragrance and were decorated with framed pictures, curtains, couches, and pillows. Participants were randomly assigned to a specific second (practice) session condition based on location (same or different) and spacing schedule (massed or spaced), with the starting room counterbalanced. Massing and spacing, same and changed contexts were between-subject manipulations. Those in the massed condition were asked to participate in two sessions, spaced 10 days apart. Those in the spaced condition were asked to participate in three sessions, the first two taking place on consecutive days and the third taking place 10 days after the second session.

Half of the participants in each spacing condition were assigned to complete the practice session in a different room from the encoding session (changed context) and return to the first room for the final test. The remaining participants completed all sessions in the same room. After reading and signing a letter of consent, participants were asked to read over the directions for the encoding session. The basic experimental procedure was similar to Cepeda et al. (2009, Exp. 1) and Simone et al. (2013) except for the "changed context" conditions. In the encoding session, participants learned 20 Swahili–English word pairs on the computer. First, the word pairs appeared in black, all-caps lettering on a gray background of a computer screen for 7 s, with Swahili words always displayed above the English translations. After all 20 word pairs had cycled through (at random), a Swahili word was shown with a text box below in which the participant was prompted to type the translation (again, order of presentation was random). If the participant typed the correct translation, the computer displayed "correct" in green lettering along with the correct answer. If the participant gave up (by pressing "Enter") or typed an incorrect response, then the computer displayed "wrong" in red lettering and then showed the correct word pair for 5 s. To complete the encoding session, participants were required to correctly produce each correct English word twice. As items were learned to a criterion of twice correct, they were dropped from the learning set, similar to Simone et al. (2013; see also Cepeda et al., 2009).

The study-phase practice session occurred either immediately after the encoding session (massed) or 24 h later (spaced). The practice session occurred either in the same room or different room as the initial encoding phase. In this session, the computer randomly presented the Swahili word prompt with an empty text box, and participants were given corrective feedback for their response, as in the encoding session. Once all word pairs had been presented, the computer randomly presented the Swahili word prompts a second time, and participants were given corrective feedback for all their responses.

In the final testing phase, which occurred 10 days after the practice session in the same room as the encoding session, participants were presented with the 20 Swahili word prompts in a random order and were asked to recall the English translations by typing them in the blank text box. Participants did not receive corrective feedback during this final test.

Results and discussion

There was no significant difference in final session performance as a function of our counterbalanced starting room ($F < 1$), which was not considered further. Our primary interest is performance data in the final test session, which took place 10 days after the practice session, but first we present data from the earlier portions of the study, which could influence final test performance (data are presented in Table 1).

Encoding session analysis

The key data from the encoding session included number of trials taken to reach criterion (recalling each English translation correctly twice). The number of trials was analyzed with a 2 × 2 × 2 factorial analysis of variance (ANOVA) with age (young vs. old), spacing schedule (massed vs. spaced), and location of the practice session (same vs. changed) as between-subjects factors. There was a significant main effect of age, $F(1, 172) = 43.8$, $p < .001$, $\eta_p^2 = .20$. There was no main effect of spacing, $F(1, 172) = 0.2$, $p = .63$, no main effect of location, $F(1, 172) = 3.3$, $p = .07$, nor were any of the interactions significant. Consistent with Simone et al. (2013), older adults took significantly more trials to reach criterion ($M = 110$, $SD = 37$) than younger adults ($M = 80$, $SD = 25$).

Table 1. Mean (*SD*) performance as a function of age, encoding condition, spacing condition, and retrieval attempt. For the encoding phase, total trials are shown. For the remaining phases, proportion recalled is shown.

	Encoding		Retrieval			
Older adult	*N*	Total trials	Part 1	Part 2	Total	Test
Changed						
Massed	21	115.7 (37.8)	0.76 (0.19)	0.81 (0.15)	0.79 (0.16)	0.35 (0.19)
Spaced	22	120.5 (41.2)	0.51 (0.19)	0.61 (0.22)	0.56 (0.20)	0.38 (0.18)
Same						
Massed	22	107.6 (32.4)	0.76 (0.18)	0.80 (0.14)	0.78 (0.15)	0.36 (0.24)
Spaced	21	97.6 (35.9)	0.63 (0.19)	0.80 (0.15)	0.72 (0.15)	0.58 (0.14)
Younger adult Changed						
Massed	26	79.5 (24.3)	0.84 (0.27)	0.88 (0.20)	0.89 (0.19)	0.57 (0.22)
Spaced	24	80.8 (18.7)	0.80 (0.14)	0.92 (0.11)	0.86 (0.12)	0.67 (0.2)
Same						
Massed	25	81.2 (31.6)	0.89 (0.14)	0.96 (0.07)	0.92 (0.10)	0.48 (0.19)
Spaced	20	76.2 (22.3)	0.77 (0.15)	0.93 (0.08)	0.85 (0.11)	0.71 (0.18)

Practice (study-phase retrieval) session

Proportion of correct responses to the Swahili prompts was calculated across all 40 responses in the practice session again with a factorial ANOVA with age (young vs. old), spacing schedule (massed vs. spaced), and location of the practice session (same vs. changed) as between-subjects factors. Note that data for one young adult participant was unavailable for this phase of the study resulting from a corrupt data file.

There was a significant main effect of age, $F(1, 172) = 55.2$, $p < .001$, $\eta_p^2 = .24$, spacing, $F(1, 172) = 18.9$, $p < .001$, $\eta_p^2 = .10$, and context, $F(1, 172) = 3.9$, $p = .048$, $\eta_p^2 = .02$. Of the two-way interactions, only age x spacing was significant, $F(1, 172) = 4.1$, $p = .044$, $\eta_p^2 = .02$ (spacing x context, $F(1, 172) = 1.9$, $p = .16$ and age x context, $F(1, 172) = 2.0$, $p = .16$ were not); the three-way interaction was also significant, $F(1, 172) = 5.2$, $p = .02$, $\eta_p^2 = .03$.

Given our primary interest in age-related effects, we conducted follow up analyses on the three-way interaction by further analyzing the practice (retrieval) data separately for both older and younger adults using a 2 × 2 ANOVA with spacing and context as factors. For older adults, there was both a significant main effect of spacing, $F(1, 82) = 15.9$, $p < .001$, $\eta_p^2 = .16$, and context, $F(1, 82) = 4.6$, $p = .04$, $\eta_p^2 = .05$, as well as a significant spacing x context interaction, $F(1, 82) = 5.3$, $p = .02$, $\eta_p^2 = .06$. We followed the interaction with independent-means *t*-tests comparing performance for all older adult participants in massed and spaced conditions for the same and changed contexts separately. Here we found no performance difference when in the same context, $t(41) = 1.3$, $p = .19$ ($M_{massed} = .78$, $SD = .15$, $N = 22$ and $M_{spaced} = .72$, $SD = .15$, $N = 21$) but there is an effect in the context change condition, $t(41) = 4.1$, $p < .001$ such that performance for older massed participants is significantly better than spaced participants ($M_{massed} = .79$, $SD = .16$, $N = 21$ and $M_{spaced} = .56$, $SD = .20$, $N = 22$). For younger adults, the comparable analysis revealed no significant effects: There was no main effect of spacing, $F(1, 90) = 3.5$, $p = .06$ or context, $F(1, 90) = 0.2$, $p = .64$ nor was the spacing x context interaction significant, $F(1, 90) = 0.5$, $p = .48$, with the overall $M = .88$, $SD = .14$, $N = 95$. Therefore, except for the older adults in the spaced context-change condition, who recalled fewer words than older adults in the massed context-change condition, performance was otherwise equalized across conditions following this practice session. In other words, study-phase retrieval is equivalent for younger adults in all conditions (spacing and context). However, there is a study-phase retrieval failure of older adults in the spaced-changed condition.

Test session (10-day retention interval)

The critical data from the test session were the proportion of correct responses to the Swahili prompts. We analyzed proportion correct data from the test session, shown in Figure 1, using a 2 × 2 × 2 ANOVA with spacing (massed or spaced), context (same or changed), and age (young or old) as between-subject factors.

There was a significant main effect of spacing, $F(1, 173) = 23.6$, $p < .001$, $\eta_p^2 = .12$, and age, $F(1, 173) = 43.6$, $p < .001$, $\eta_p^2 = .20$, but no main effect of context, $F(1, 173) = 1.8$, $p = .18$. Of the interactions, spacing x context, $F(1, 173) = 7.4$, $p = .007$, $\eta_p^2 = .04$, and context x age, $F(1, 173) = 4.8$, $p = .03$, $\eta_p^2 = .03$ were significant. The remaining two

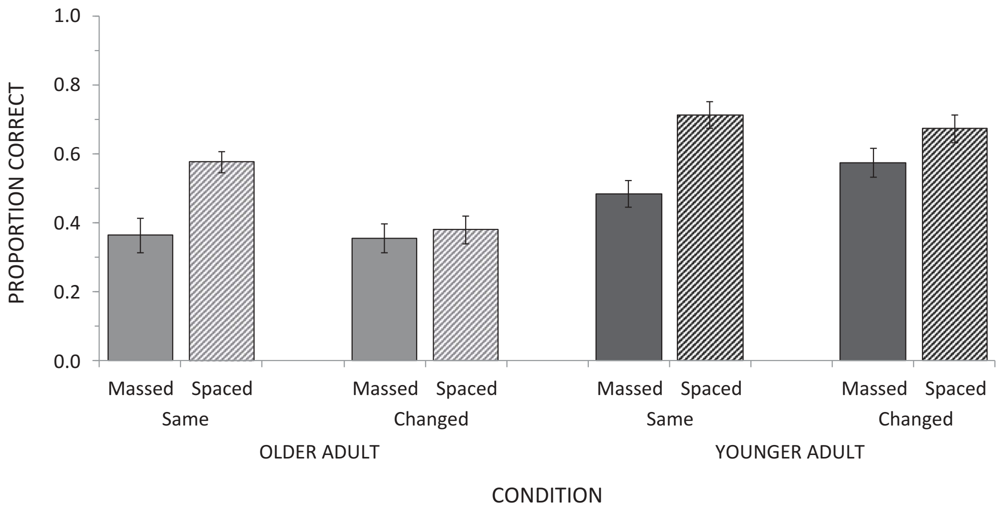

Figure 1. Proportion overall correct (error bars represent ±*SEM*) in the Test Session as a function of spacing (massed or spaced) and environmental context (same or changed room) for older and younger adults.

interactions were not significant: spacing x age, $F(1, 173) = 0.06$, $p = .44$, and spacing x context x age, $F(1, 173) = 0.26$, $p = .61$. We evaluated each of the interactions with follow up *t*-tests.

For the spacing x context interaction, we analyzed spacing effects separately for each context (collapsing across age). In the same context, there was a significant spacing effect, $t(86) = 5.1$, $p < .001$, with better performance for participants in the spaced condition ($M = .64$, $SD = .17$, $N = 41$) than in the massed condition ($M = .43$, $SD = .22$, $N = 47$). In the changed context, there was no significant benefit to spacing, $t(91) = 1.2$, $p = .25$, with participants in the spacing condition showing comparable performance ($M = .53$, $SD = .24$, $N = 46$) to participants in the massed condition ($M = .47$, $SD = .23$, $N = 47$). However, as Figure 1 shows, although the interaction with age was not significant, the trend was that context change caused a decrease in recall in the spaced condition for older adults but context change caused an increase in recall in the massed condition for younger adults.

For the context x age interaction, we analyzed context effects separately for each age group. For older adults, there was a significant cost to being in the changed context, $t(84) = 2.3$, $p = .03$ ($M = .37$, $SD = .19$, $N = 43$), compared to being in the same context ($M = .46$, $SD = .22$, $N = 43$), regardless of spacing condition. Context change did not affect younger adult performance, $t(93) = 0.8$, $p = .42$, ($M_{changed} = .62$, $SD = .21$, $N = 50$ and $M_{same} = .58$, $SD = .21$, $N = 45$).

General discussion

Our goal was to investigate the effect of contextual changes on the spacing benefit in older adults (in comparison to performance of younger adults) and to evaluate the role of context in conferring a spacing benefit. We found that both younger and older adults benefited from spacing when all sessions were in the same room, consistent with previous work (Balota et al., 1989; Kornell, Castel, Eich, & Bjork, 2010; Simone et al., 2013). However, the spacing benefit was eliminated in both older and younger adults in the final test when practice occurred in a different context from initial learning (i.e., a different room). Since the change in context affected recall of older adults in both the practice session and at the test and only at the test for younger adults, different processes could be at work for the two age groups.

Considering the context-same condition first, examination of the data in Figure 1 suggest that when the context was consistent between initial learning and the study-phase practice session, both older and younger adults benefited from a spacing schedule, comparable to the findings of Simone et al. (2013). Simone et al. reported that those who experienced a spacing gap 24 h after initial encoding of Swahili–English word pairs remembered significantly more over a 10-day retention interval than those who studied immediately after initial encoding. Our results are generally consistent with these findings, in that spacing led to improved recall. Unlike Simone et al., however, we did not find an age x spacing interaction. Their findings suggest that older adults benefited less from spacing than younger adults (they reported a η_p^2 of .061, suggesting a medium effect size) whereas we found that, when context was constant, it benefited both groups equally, although younger adults overall recalled more words than older adults. The source of this difference in these findings is unknown, particularly given our same-context procedure was a direct replication of that experiment. One possibility we investigated was that there could have been differences in the encoding session. Given the relatively small difference in mean trials to criterion (M = 110, SD = 37 trials to criterion for older adults in the present study and M = 136, SD = 62 in Simone et al.), we believe this is an unlikely explanation. Even so, our lack of an age x spacing interaction is consistent with other studies (Balota et al., 1989; Kornell et al., 2010).

Examination of Retrieval 1 (Table 1) suggests that older adults benefitted from consistent contextual cues when there was a 1-day delay between encoding and retrieval. In other words, context matters for older adults when it was available to cue recall. This context effect is not evident for young adults – they appear minimally sensitive to contextual cues even after a 1-day delay (although this could be the result of a ceiling effect). Following the 10-day retention interval, context became less relevant for older adults. In other words, degree of retrieval success during the second session appears to have affected performance in the test session (i.e., a testing effect; Roediger & Karpicke, 2006).

Spacing effect and context

One key difference between our study and others (Balota et al., 1989; Kornell et al., 2010; Simone et al., 2013) is that we explicitly manipulated the context of the practice sessions to determine the effect of environmental context manipulation on the spacing benefit in younger and older adults. Using different methodologies in spacing studies with younger adults, others have found that a change in context during spacing studies improves memory in younger adults. For example, Smith et al. (1978) found that spacing in a different room resulted in greatest delayed recall. But Smith et al. procedure differed from the one reported here in several ways, including the spacing gap (3 h vs. 24 h), retention interval (3 h vs. 10 days), and testing room (novel room vs. original encoding room). Similarly, Verkoeijen et al. (2004) found that a change in context improved massed performance and impaired spaced learning, a result supporting both contextual variability theory and study-phase retrieval hypothesis of the spacing benefit. Their procedure also differed from ours in study type (item-level rather than list-level), context changed (background screen images changing rather than room changes), recall (free vs. cued), and study design (within vs. between).

We found that contextual change during the study-phase practice session resulted in the elimination of the spacing benefit relative to performance in massed conditions in both older and younger adults. Our results are consistent with contextual variability theory of spacing (e.g., Madigan, 1969; Melton, 1970; also the related Estes, 1955; stimulus-fluctuation model) in that changing context during the practice session impacted performance at the final test in that neither younger or older adults showed a benefit to spacing in the final recall. The effect of context on performance did not interact with age; however, there was a trend for younger adults to show improved massed performance (while not significant, it is in the same direction as what was reported by Verkoeijen et al., 2004) and older adults to show diminished spaced performance under context-change conditions, perhaps due to differences in processing contextual information.

Age and context

There is evidence suggesting that older adults rely differently on contextual cues compared to younger adults, with some theories suggesting over-reliance on contextual cues to compensate for lack of other encoding strategies, (Craik, 1983, 1986), less ability to utilize contextual cues (Light & Singh, 1987), and age-related decrease in sampling rates of contextual information (Balota et al., 1989). Although criterion learning equated performance on one key dimension following initial learning, at the end of the study-phase practice session, which included two cued recall trials (with corrective feedback), there was an age-related difference in performance with younger adults recalling more cued words than older adults (a finding consistent with Balota et al.). Additionally, our study showed that the combination of a change in context and a 1-day delay had an immediate negative impact on older adults' performance in the study-phase practice session of the study, a finding that supports the study-phase retrieval hypothesis of spacing such that the spacing benefit is reliant on successful retrieval of the information in the study phase.

Unlike the performance of older adults, contextual change during the practice session did not affect performance of younger adults in the spaced conditions, who recalled 85% (same context) and 86% (changed context) of the word pairs correctly in the practice session, rates comparable to the massed conditions (92% for the same context and 89% for the changed context). Thus, the spacing gap and change in context did not, to the same extent, negatively impact recall of younger adults. It appears older adults relied more heavily on consistent contextual cues in this spaced practice session and thus recalled fewer word pairs when cued for retrieval in a different room after a delay.

This deficit in recall after a delay and room change also suggests that, compared to younger adults, older adults were more reliant on consistent environmental cues to help cue study-phase retrieval, a finding that is consistent with Craik's (1983, 1986) environmental support hypothesis. Following a 24-h spacing delay, under the context-same condition, older adults correctly recalled 72% of the word pairs, whereas those in the context-change condition recalled only 56% of the word pairs.

Campbell et al. (2014) found that older adults are more apt to bind irrelevant information to target information than younger adults, which is the dark side of contextual cuing. If older adults rely heavily on consistent contextual cues, as we have found here and is in line with Craik's environmental hypothesis, then rather than learn to associate the two words, older adults are binding irrelevant but in some circumstances helpful contextual cues to the target word pairs. Here the binding or over-reliance on the environment became more of a distraction than a benefit to learning, even though recall was cued – the most important cue, the word that cued recall, was constant. To the extent that the cue prompts recall, additional contextual cues should be irrelevant.

We have shown that the spacing benefit is sensitive to what occurs during the study-phase practice session such that if the practice occurs in a novel environment from initial encoding, then the spacing benefit is eliminated in both younger and older adults. Another place to look for why this spacing benefit difference exists is to examine performance during the encoding session for clues. Similar to Simone et al. (2013), older adults experienced more repetitions and took longer to learn the word pairs in the encoding session. However, because we used a criterion learning procedure, all participants were functionally equal at least in terms of having been successful twice for each word pair. There is some evidence (e.g., Trahan & Larrabee, 1992) to suggest that differences in performance at test may be a result of differences at encoding. We examined this by looking at number of trials to criterion and test performance and found a negative correlation ($r(180) = -.53$, $p < 001$). This held true for both younger and older adults. Those who took longest to reach criterion recalled the fewest words at Test, regardless of their spacing condition or context manipulation. This is unrelated to our experimental manipulation but it does speak to the importance of initial encoding to later retrieval and suggests that this spacing schedule (24-h gap and 10-day retention interval) was not ideal for slower learners, or possibly that slower learners do not benefit from spacing.

Another possible cause of the difference in test performance is because older adults took longer to reach criterion in the first session they had more exposure to the word pairs. When considering an item-level study, that could be problematic: words learned first were more likely to be spaced and words learned later were more likely to be massed within the encoding session. However, in a list-level study such as reported here, even the last learned items were initially spaced, at least through the early part of the encoding session. How these repeated exposures of both spaced (early trials) and massed (later trials) for the later learned word pairs might interact in a list-level study is not clear and is beyond the scope of our study.

Differences between our study and others may also prove important. For example, Kornell et al. (2010) used paintings and artist names, Luo et al. (2007) presented words and pictures, and Balota et al. (1989) presented English noun word pairs. These stimulus differences may have different implications for learning and memory, including differences in performance for older and younger adults. Furthermore, additional methodological differences exist between the study reported here and previous work (e.g., cued recall, significant spacing gap, and retention interval), making any direct comparisons difficult.

The role of context in learning is a complex one. Contextual cues contribute to the spacing benefit, although not all data can be explained by a purely contextual variation explanation (e.g., Ross & Landauer, 1978). However, it appears that a contextual component may be necessary for a more predictive model of the mechanism, based on findings that a non-contextual model (Pavlik & Anderson, 2005) cannot explain observed contracting and expanding spacing interval effects (Küpper-Tetzel, Kapler, & Wiseheart, 2014; based on modeling by Lindsey, Mozer, Cepeda, & Pashler, 2009). However, contextual processing is an assumption of both contextual variability theory and study-phase retrieval hypothesis and, as was found in other studies using different stimuli and experimental design, our study also showed that explicit contextual changes can affect the spacing benefit in younger and older adults.

In conclusion, we found a spacing benefit for both younger and older adults, but that benefit disappeared for both age groups when the practice session was in a different location. Furthermore, we found that the context change in the practice session appears most detrimental to older adults' recall and relearning of the word pairs. Not only was older adults' performance lower in the practice (study-phase retrieval) session, performance did not recover. Practically speaking, our findings suggest that older adults would benefit most from spaced retrieval of to-be-remembered associated items (such as face-name pairs) in the same context in which they learned the association.

Acknowledgments

This research was supported by the Gerald and Sally DeNardo Scholarship and the Doelger Fellowship granted to KEB. Thank you to Nicole Assumpçao, Paul David, Ashleigh MacLean, Julia Papanek, Gabriel Salazar, and Jake Teeny for assistance in data collection and to the anonymous reviewers for their constructive feedback. KEB is now at Harvard University.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Gerald and Sally DeNardo Scholarship; Doelger Fellowship.

ORCID

Matthew C. Bell http://orcid.org/0000-0002-0241-8035
Melody Wiseheart http://orcid.org/0000-0001-8332-6775

References

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

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. doi:10.1037/0882-7974.4.1.3

Bell, M. C., Kawadri, N., Simone, P. M., & Wiseheart, M. (2014). Long-term memory, sleep, and the spacing effect. *Memory*, *22*, 276–283. doi:10.1080/09658211.2013.778294

Campbell, K. L., Trelle, A., & Hasher, L. (2014). Hyper-binding across time: Age differences in the effect of temporal proximity on paired-associate learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 293–299. doi:10.1037/a0034109

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology*, *56*, 236–246. doi:10.1027/1618-3169.56.4.236

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380. doi:10.1037/0033-2909.132.3.354

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *302*, 341–359. doi:10.1098/rstb.1983.0059

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

Craik, F. M., & Rose, N. S. (2011). Memory encoding and aging: A neurocognitive perspective. *Neuroscience & Biobehavioral Reviews*, *36*, 1729–1739. doi:10.1016/j.neubiorev.2011.11.007

Delaney, P. F., Verkoeijen, P. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In B. H. Ross & B. H. Ross (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 53, pp. 63–147). San Diego, CA: Elsevier Academic Press. doi:10.1016/S0079-7421(10)53003-2

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*, 795–805. doi:10.1037/0021-9010.84.5.795

Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review*, *62*, 369–377. doi:10.1037/h0046888

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, *7*, 95–112. doi:10.3758/BF03197590

Glenberg, A. M., & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory & Cognition*, *8*, 528–538. doi:10.3758/BF03213772

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*, 325–331. doi:10.1111/j.2044-8295.1975.tb01468.x

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

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, *25*, 498–503. doi:10.1037/a0017807

Küpper-Tetzel, C. E. (2014). Understanding the distributed practice effect: Strong effects on weak theoretical grounds. *Zeitschrift Für Psychologie*, *222*, 71–81. doi:10.1027/2151-2604/a000168

Küpper-Tetzel, C. E., Kapler, I. V., & Wiseheart, M. (2014). Contracting, equal, and expanding learning schedules: The optimal distribution of learning sessions depends on retention interval. *Memory & Cognition*, *42*, 729–741. doi:10.3758/s13421-014-0394-1

Light, L. L., & Singh, A. (1987). Implicit and explicit memory in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 531–541. doi:10.1037/0278-7393.13.4.531

Lindsey, R., Mozer, M. C., Cepeda, N. J., & Pashler, H. (2009). Optimizing memory retention with cognitive models. In A. Howes, D. Peebles, & R. Cooper (Eds.), *Proceedings of the ninth international conference on cognitive modeling (ICCM 2009)* (pp. 74–79). Manchester, UK: ICCM.

Luo, L., Hendriks, T., & Craik, F. I. M. (2007). Age differences in recollection: Three patterns of enhanced encoding. *Psychology and Aging*, *22*, 269–280. doi:10.1037/0882-7974.22.2.269

Maddox, G. B., Balota, D. A., Coane, J. H., & Duchek, J. M. (2011). The role of forgetting rate in producing a benefit of expanded over equal spaced retrieval in young and older adults. *Psychology and Aging*, *26*, 661–670. doi:10.1037/a0022942

Madigan, S. A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 828–835. doi:10.1016/S0022-5371(69)80050-2

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 596–606. doi:10.1016/S0022-5371(70)80107-4

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, And Cognition*, *26*, 1170–1187. doi:10.1037/0278-7393.26.5.1170

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*, 104–118. doi:10.1037/0882-7974.23.1.104

Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, *29*, 559–586. doi:10.1207/s15516709cog0000_14

Raaijmakers, J. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, *27*, 431–452. doi:10.1016/S0364-0213(03)00007-7

Roediger, H L., & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x

Ross, B. H., & Landauer, T. K. (1978). Memory for at least one of two items: Test and failure of several theories of spacing effects. *Journal of Verbal Learning & Verbal Behavior*, *17*, 669–680. doi:10.1016/S0022-5371(78)90403-6

Simone, P. M., Bell, M., & Cepeda, N. J. (2013). Diminished but not forgotten: Effects of aging on magnitude of spacing effect benefits. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *68*, 674–680. doi:10.1093/geronb/gbs096

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, *6*, 342–353. doi:10.3758/BF03197465

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

Trahan, D. E., & Larrabee, G. J. (1992). Effect of normal aging on rate of forgetting. *Neuropsychology*, *6*, 115–122. doi:10.1037/0894-4105.6.2.115

Verkoeijen, P. L., Rikers, R. P., & Schmidt, H. G. (2004). Detrimental influence of contextual change on spacing effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 796–800. doi:10.1037/0278-7393.30.4.796