Title
Investigating the Effects of Spacing on Working Memory Training Outcome: A Randomized, Controlled, Multisite Trial in Older Adults.

Journal
Journal of Gerontology Series B: Psychological Sciences and Social Sciences, 75(6)

Authors
Jaeggi, Susanne
Buschkuehl, Martin
Parlett-Pelleriti, Chelsea
et al.

Publication Date
2020-06-02

DOI
10.1093/geronb/gbz090

Peer reviewed

Journals of Gerontology: Psychological Sciences
cite as: J Gerontol B Psychol Sci Soc Sci, 2020, Vol. 75, No. 6, 1181–1192
doi:10.1093/geronb/gbz090
Advance Access publication July 29, 2019

Special Section: Cognitive Interventions in Late Life

Investigating the Effects of Spacing on Working Memory Training Outcome: A Randomized, Controlled, Multisite Trial in Older Adults

Susanne M. Jaeggi, PhD,[1,*] Martin Buschkuehl, PhD,[2] Chelsea M. Parlett-Pelleriti, MS,[3] Seung Min Moon, BS, BA,[1] Michelle Evans, BA,[4] Alexandra Kritzmacher, MA,[4] Patricia A. Reuter-Lorenz, PhD,[4] Priti Shah, PhD,[4] and John Jonides, PhD[4]

[1]School of Education, University of California, Irvine. [2]MIND Research Institute, Irvine, California. [3]Schmid College of Science and Technology, Chapman University, Orange, California. [4]Department of Psychology, University of Michigan, Ann Arbor.

*Address correspondence to: Susanne M. Jaeggi, PhD, School of Education and School of Social Sciences (Department of Cognitive Sciences), University of California, Irvine, CA. E-mail: smjaeggi@uci.edu

Received: January 3, 2019; Editorial Decision Date: July 11, 2019

Decision Editor: Brent Small, PhD

Abstract

Objective: The majority of the population will experience some cognitive decline with age. Therefore, the development of effective interventions to mitigate age-related decline is critical for older adults' cognitive functioning and their quality of life.

Methods: In our randomized controlled multisite trial, we target participants' working memory (WM) skills, and in addition, we focus on the intervention's optimal scheduling in order to test whether and how the distribution of training sessions might affect task learning, and ultimately, transfer. Healthy older adults completed an intervention targeting either WM or general knowledge twice per day, once per day, or once every-other-day. Before and after the intervention and 3 months after training completion, participants were tested in a variety of cognitive domains, including those representing functioning in everyday life.

Results: In contrast to our hypotheses, spacing seems to affect learning only minimally. We did observe some transfer effects, especially within the targeted cognitive domain (WM and inhibition/interference), which remained stable at the 3-month follow-up.

Discussion: Our findings have practical implications by showing that the variation in training schedule, at least within the range used here, does not seem to be a crucial element for training benefits.

Keywords: Cognitive training, Distributed learning, Transfer

Although only a minority of adults aged more than 65 years will develop Alzheimer's disease, the vast majority will experience some decline in cognitive function with age. Although people vary greatly in the extent of cognitive decline, even subtle losses can substantially affect everyday life, adversely affecting critical decisions about health care, retirement, and other issues faced daily by millions of older adults (Alzheimer's Association, 2015; Tomaszewski Farias et al., 2009). Consequently, it is critically important to develop effective interventions that can contribute to the prevention of cognitive decline or improve older adults' cognitive functions, and ultimately, contribute to their well-being and quality of life.

There is growing evidence that some cognitive interventions may be effective in facilitating a range of cognitive skills with improvements evident even in healthy older adults and those with mild cognitive impairment (cf. Hill et al., 2017; Karbach & Verhaeghen, 2014; Weicker, Villringer, & Thöne-Otto, 2016 for recent meta-analyses). Some of these effects have been shown to be very long lasting (e.g., Rebok et al., 2014). Although some controversy surrounds their real-world benefits (e.g., Simons et al., 2016), the feature shared by many promising interventions is the training focus on basic cognitive skills, often related to working memory (WM; Karbach & Verhaeghen, 2014; Lustig, Shah, Seidler, & Reuter-Lorenz, 2009). WM is the cognitive system supporting the active maintenance of task-relevant information during the performance of a cognitive task; it is one of the main mechanisms facilitating purposeful behavior (Shah & Miyake, 1999). WM underlies performance in virtually all complex cognitive tasks and critically, it is highly susceptible to age-related decline (Bugg, Zook, DeLosh, Davalos, & Davis, 2006; Park et al., 2002) and predictive for everyday functioning in old age (Cahn-Weiner et al., 2007; Tomaszewski Farias et al., 2009). Furthermore, experimental and neuroimaging studies have demonstrated a close relationship between WM and episodic memory (e.g., Buckner, 2004; Flegal & Reuter-Lorenz, 2014). Corroborating those findings, several intervention studies observed improvements in various measures of episodic memory after WM training (e.g., Buschkuehl et al., 2008; Flegal, Ragland, & Ranganath, 2019; Richmond, Morrison, Chein, & Olson, 2011), which might be driven by common neural networks between the two cognitive domains (Dahlin, Neely, Larsson, Backman, & Nyberg, 2008).

Despite promising demonstrations that WM training is effective in older adults (e.g., Borella, Carretti, Zanoni, Zavagnin, & De Beni, 2013; Stepankova et al., 2014), several studies report only minimal transfer (e.g., Brehmer et al., 2011; Richmond et al., 2011), whereas others report effects in young adults with more limited or no effects in older adults using the same intervention (e.g., Brehmer, Westerberg, & Bäckman, 2012; Dahlin et al., 2008). Such findings suggest that some current interventions may not meet the needs of older adults, that age-related limitations in plasticity are difficult to overcome, or both. As such, more research is needed to make WM interventions more suitable for older adults, and also, to make them more robust by determining training features that might moderate transfer in an aging population. Specifically, systematic research is needed to determine the optimal scheduling of the training sessions, that is, the effect of spacing to maximize learning outcomes. Despite the extensive literature on spaced versus massed training in the domains of skill and verbal learning and their enormous practical significance (e.g., Cepeda et al., 2009; Krug, Davis, & Glover, 1990), spacing effects have been rarely tested in the context of WM training. To the best of our knowledge, there is one study to date demonstrating that spaced and massed practice produced minimal differences in specific training effects, yet distributed practice had a significant impact on immediate transfer (post-test performance). Specifically, the group showing the largest transfer effects completed one training session per day (as opposed to the groups that trained multiple times per day; Wang, Zhou, & Shah, 2014), which is consistent with a host of prior results demonstrating that distributed practice leads to better learning and retention than massed practice (Cepeda et al., 2009). Because this study was conducted in children, it is unclear whether and how the results might translate to an older population.

To address these outstanding questions in the literature, this study investigates the spacing of training sessions and how different spacing schedules might affect training performance and transfer in an older adult population. We focus on long-term outcomes given that the effects of spacing have often been observed after a longer delay (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). The primary goal of this study was to test whether or not the training regimen affected WM and related processes such as inhibitory control. Given previous reports, we also tested whether training WM might transfer to measures of episodic memory, especially those that serve as proxy for everyday memory functions and problem solving (Cantarella, Borella, Carretti, Kliegel, & de Beni, 2017; Weicker et al., 2016). Specifically, we conducted a randomized controlled multisite trial, in which a relatively large sample of older adult participants was assigned to either an intervention targeting WM, or an alternative intervention that focused on general knowledge and vocabulary learning. All participants were instructed to complete 20 sessions of training at home using a tablet device. Within both interventions, participants were further assigned to complete their sessions twice per day, once per day, or once every-other-day. Before and after the intervention, as well as 3 months after training completion, participants' cognitive functions in various domains were tested in order to assess potential transfer effects.

On the basis of the previous literature, we predicted that the WM-training group would outperform the knowledge-training group in measures closely related to the trained task, specifically, in WM and interference/inhibitory control. We also explored potential improvements in long-term memory (LTM) measures, especially those that are proxies for real-world performance. We did not expect improvements in basic processing speed and vocabulary, which served as control measures; if anything, the knowledge-training group should show more pronounced improvement in the vocabulary measure as compared to the WM-training group given that this was the focus of the alternative intervention.

Within the WM-training group, we predicted that the every-other-day group would outperform the twice-per-day group (and potentially, the every-day group) during training, and especially during the long-term follow-up sessions for which we were expecting the effects of spacing to become most apparent (Cepeda et al., 2006). We also expected the effects of spacing to be most pronounced in the measures that are most closely related to the trained task.

Method

Participants and Design

After passing an initial phone screening to establish eligibility criteria, 183 participants were randomly assigned to either the WM- or knowledge-training group. Furthermore, within each group, participants were randomly assigned to one of three spacing conditions that required them to complete their intervention twice per day, once per day, or once every-other-day (see Figure 1 and Supplementary Figure 1). After intervention completion, participants returned for the post-test, as well as for a follow-up assessment 3 months later. Fifty-seven percent of the participants were recruited in Southern California, and 43% were recruited in Southeast Michigan. All research procedures were approved by an institutional review board and participants signed an informed consent.

Demographic variables for the intervention groups, and for those who did not complete the study or were excluded, as well as further participant details are provided in the Supplementary Material, p. 2, and Supplementary Table 1. The two intervention groups did not differ significantly in any of the demographic variables. Importantly, the individuals who were excluded or dropped out of the interventions were not different from the analyzed sample, except for self-reported socioeconomic status (SES; the analyzed sample reported lower SES), and the number of completed training sessions (the analyzed sample completed more sessions; see Supplementary Table 1).

Tasks and Materials

Training tasks

Working memory training. Participants trained on a tablet-based version of the n-back task that used pictures as stimuli, and required indicating whether a presented picture was the same as the one presented n trials previously (similar to the task used as outcome measure in Katz, Jaeggi, Buschkuehl, Stegman, & Shah, 2014). The stimuli were presented in a moving window that lasted for 1,000 ms with an interstimulus interval of 2,500 ms. Each stimulus could be a target trial, a nontarget trial, or a lure trial (i.e., stimuli that were the same as the target stimuli, except that they were presented in the wrong position, such as n ± 1 back). The task was adaptive in that the level of difficulty (number of n − 1/n + 1 lures and n-back level) changed with respect to participants' performance after each round (Jaeggi, Buschkuehl, Shah, & Jonides, 2014). Participants worked on the same n-back level with no lures, few lures (i.e., two), and many lures (i.e., six), and moved up an n-back level after clearing the many-lures round or moved down an n-back level after making too many errors in the no-lure round. Participants completed 10 rounds per training session, and each round consisted of five target trials, 10 + n nontargets trials, a variable number of lures (zero, two, or six), and one filler trial at the start of the

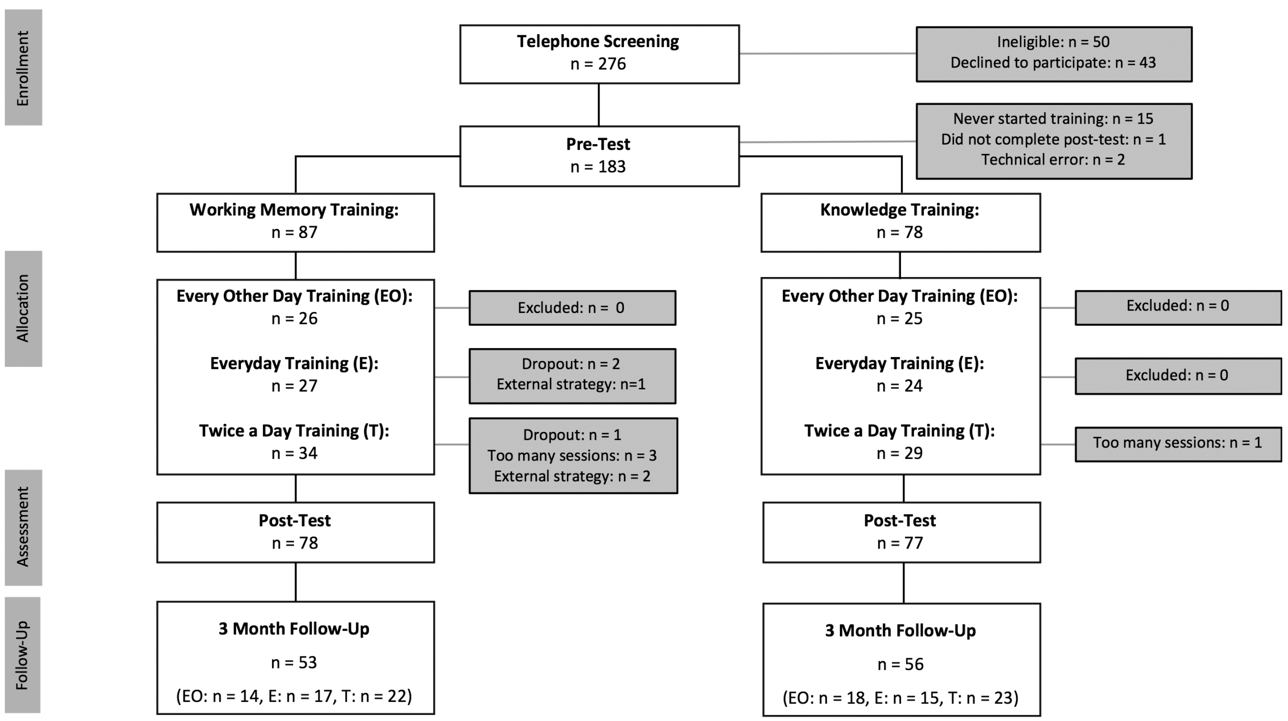

Figure 1. Flowchart of participants and allocation.

round. The dependent variable was the average level of n reached per training session.

Knowledge skills training. Participants used a tablet-based general knowledge task as described previously (Jaeggi et al., 2014). Participants were presented with general knowledge and vocabulary questions, along with four answer alternatives. After participants selected their response, they received feedback, and questions answered incorrectly were presented again in the beginning of the next session in order to promote learning. The task was adaptive in that the difficulty of the questions was reflected in levels as determined by pilot testing. Although the emphasis was on accuracy, participants were given a time limit of 45 s to respond before a trial was automatically marked as incorrect. The dependent variable was the average level of difficulty achieved in each training session.

Baseline/control assessments

We administered a detailed *demographic and health questionnaire* to assess participants' general physical and psychological health and well-being, including the *WHOQOL-OLD* (Fang et al., 2012). We also screened for general cognitive status using the *Mini-Mental State Examination* (Tombaugh & McIntyre, 1992), *Geriatric Depression Scale* (Yesavage, 1988), and *Generalized Anxiety Depression Questionnaire* (Spitzer, Kroenke, Williams, & Löwe, 2006; cf. Supplementary Table 1).

The following measures were administered at baseline, as well as during post- and follow-up sessions, although they were considered control measures, for which we did not expect any changes.

Vocabulary

To assess general intellectual ability, we used the *Mill Hill Vocabulary test* (Raven, Raven, & Court, 1998) in which 33 target words were presented with 6 potential synonyms, and participants had to select the appropriate word that matches the target word. The dependent variable was the number of correctly identified synonyms. We used parallel-test versions counterbalanced across participants and sessions.

Processing speed

Pattern and letter comparison (Ribaupierre & Lecerf, 2006). We administered two variants of paper–pencil perceptual and motor speed tasks in which participants had to indicate as quickly as possible whether two patterns or letter strings presented next to each other are the same or not (e.g., QLXVST __ QLNSVT). There were 60 items for the pattern comparison, and 42 items for the letter comparison, and we used the total time to complete each of the tasks (in seconds) as the dependent variable.

Transfer measures

In order to test for *training-related changes* (*i.e., transfer*), we administered the following set of tasks at baseline, as well as during the post- and follow-up test.

Working memory

Spatial n-back. As a measure of near transfer, we administered a spatial version of the trained n-back task, also via tablet devices. The task parameters were the same as for the training task except that the task was not adaptive; that is, after one round of 1-back, the task difficulty was set to a 2-back level. Participants completed three rounds of 2-back without lures and three rounds of 2-back with many (six) lures. The dependent variable was pr (proportion hits minus false alarms) as well as reaction time (median) for correct trials, averaged across all 2-back trials. In addition, we used the false alarm rate separately to specifically capture aspects of interference resolution and memory precision.

Sternberg task. We used a similar version of a computerized item recognition task as used in Iordan and colleagues (2018). Participants were required to encode and retain a set of consonant letters in uppercase (set size 4–8) for several seconds, and after a brief retention interval, they were given a probe letter in lowercase and were asked to indicate whether or not this probe letter was part of the initial memory set. Participants completed three blocks of 20 trials each (4 trials per set size), and the dependent variables were accuracy and reaction time (median) for correct responses across all trials.

Symmetry span. This computerized task was adapted from Redick and colleagues (2012). Participants had to indicate whether or not a pattern was symmetrical, after which a square was presented in 1 of 16 locations on a grid. After two to six trials (symmetry decision + location), participants were asked to recall the locations in order using the computer mouse. We used parallel-test versions counterbalanced across participants and sessions, and the number of correctly recalled sets served as the dependent variable.

Inhibitory control/interference

D2 (Brickenkamp, 2002). We administered a paper–pencil version of the D2 which consists of 14 lines of letters (either p or d) with one to four dashes below and/or above each letter. Participants are provided 20 s per line and asked to cross out any ds with two dashes as quickly as possible while ignoring all other items. The total number of items completed minus any type of error (commission or omission; $TN - E$) was our index of inhibitory control.

Episodic memory

Visual LTM. This paper–pencil task was adapted from Perrig and colleagues (2006). Participants were shown two arrays of Snodgrass and Vanderwart-like line drawings of objects, patterns, and words on a single page, and they had to mark as many differences as possible between the two arrays within 3 min. About 20 min later, participants recalled all they could remember from the picture, as well as the differences they found. The dependent variables were the number of correctly recalled items. In addition, in order to capture another aspect of interference and memory precision, we used intrusions (incorrectly recalled items). We used parallel-test versions counterbalanced across participants and sessions.

Metamemory. This task was adapted from McGillivray and Castel (2011) (cf. Parlett-Pelleriti et al., 2019). Participants were presented with five 12-word lists, and after each word, they were asked to place a bet between 0 and 10 points representing the likelihood of remembering that word later on. At the end of each list, participants were asked to recall as many words as possible. For each correctly remembered word, the bet for that word was added to their score, and for failure to recall a word, it was subtracted from their score. After each list, participants were presented with their score before proceeding to the next list. Parallel-test versions were administered, counterbalanced across participants and sessions. The number of correctly recalled words across all lists served as the dependent variable.

Characterization of the Elderly on Daily Activities in the Real-World (CEDAR). This verbal prompting task was adapted from Thomas (2015) to assess the performance of everyday cognition akin to the Observed Tasks of Daily Living (Diehl et al., 2005). Participants were assigned the role of a neighbor for a fictitious character and asked by a fictitious relative to complete a series of errands that involved managing medications, meal planning, finances, and making long-term decisions (e.g., bank/doctor selection). A parallel version was used, counterbalanced across participants and sessions. We used accuracy (standardized across subtasks and averaged into one measure) as the dependent variable.

Analytical approach

Using SPSS 24, JASP 0.9.1.0, and R, we conducted analyses to investigate the effects of spacing on training performance, and potentially, generalizing effects to non-trained outcome measures (cf. Supplementary Material for further details). Our hypotheses and analytical approach have been preregistered (cf. AsPredicted #7897; https://aspredicted.org/mp2jv.pdf).

Results

Specific Training Effects

The performance of the WM-training group improved over the course of training, from an average n-back level of 2.36 (SD = 0.41) obtained across the first two sessions to an average of 2.81 (SD = 0.74) in the last two sessions ($t(77) = 6.29$, $p < .001$; $d = .72$; BF_{10} = 1.342e+6).

However, the spacing condition had no effect on training, as indicated by a repeated measures analysis of variance (ANOVA) with spacing condition as the between-subject factor and session number (1–20) as the within-subject factor (main effect of spacing condition: $F(2,58) = 0.52$, $p = .60$; $\eta^2_p = .02$; $BF_{10} = 0.29$; Spacing × Session interaction: $F(38,1102) = 0.90$, $p = .64$; $\eta^2_p = .03$; $BF_{10} = 0.001$), or with a univariate ANOVA using gain (calculated as the difference between the performance during the two last sessions and the first two sessions) as the dependent variable ($F(2,75) = 0.24$, $p = .79$; $\eta^2_p = .006$; $BF_{10} = 0.13$) (cf. Figure 2A and Table 1).

The performance of the knowledge-training group also improved, from an average difficulty level of 5.33 (SD = 1.27) across the first two sessions to an average of 7.01 (SD = 2.08) across the last two sessions ($t(76) = 12.32$, $p < .001$; $d = 1.41$; BF_{10} = 1.198e+17). Again, spacing condition had no effects on training (repeated measures ANOVA: main effect of spacing: $F(2,59) = 2.10$, $p = .13$; $\eta^2_p = .07$; $BF_{10} = 0.87$; Spacing × Session interaction: $F(38,1121) = 1.10$, $p = .32$; $\eta^2_p = .04$; $BF_{10} = 0.006$; univariate ANOVA (gain): $F(2,74) = 0.13$, $p = .88$; $\eta^2_p = .003$; $BF_{10} = 0.12$) (cf. Figure 2B and Table 1).

However, across both groups, there was a difference in the number of completed training sessions in that the twice-per-day group completed significantly more training sessions on average than the every-other-day group (M = 21.45 [SD = 3.63] versus M = 19.90 [SD = 2.03]; $t(104) = 2.56$, $p = .007$; $d = .56$; $BF_{10} = 4.76$).

To investigate differential training effects as a function of spacing conditions, we calculated univariate ANOVAs separately for each intervention group using participants' final performance (average last two sessions) as well as the gain as the dependent variables (cf. Table 1). In addition, we conducted individual curve-fitting analyses using linear as well as spline regression approaches to analyze early versus late learning (Iordan et al., 2018; Lövdén, Bäckman, Lindenberger, Schaefer, & Schmiedek, 2010) (cf. Supplementary Material). These analyses also failed to reveal any group differences due to spacing condition. The only hint of a spacing effect was a main effect of spacing condition in the change from early to late learning for the WM-training group: for the every-day group the improvement was more pronounced compared to the every-other-day group ($p = .03$; $d = .67$; $BF_{10} = 2.42$). However, no group differences were evident in any of the other outcome measures, indicating that overall, spacing effects on training were negligible for both the WM- and the knowledge-training group.

Note that there was some variability in adherence to the prescribed training schedule for various reasons (e.g., forgetting to train, travel, illness, catching up with missed training sessions). Thus, we conducted additional analyses on only those participants who strictly followed the assigned schedule across all 20 training sessions. Once again, no effects of spacing emerged for any of the training measures (see Figure 3A and B, and Supplementary Table 2).

Transfer Effects

Descriptive information for each of the measures as a function of group and testing session is provided in

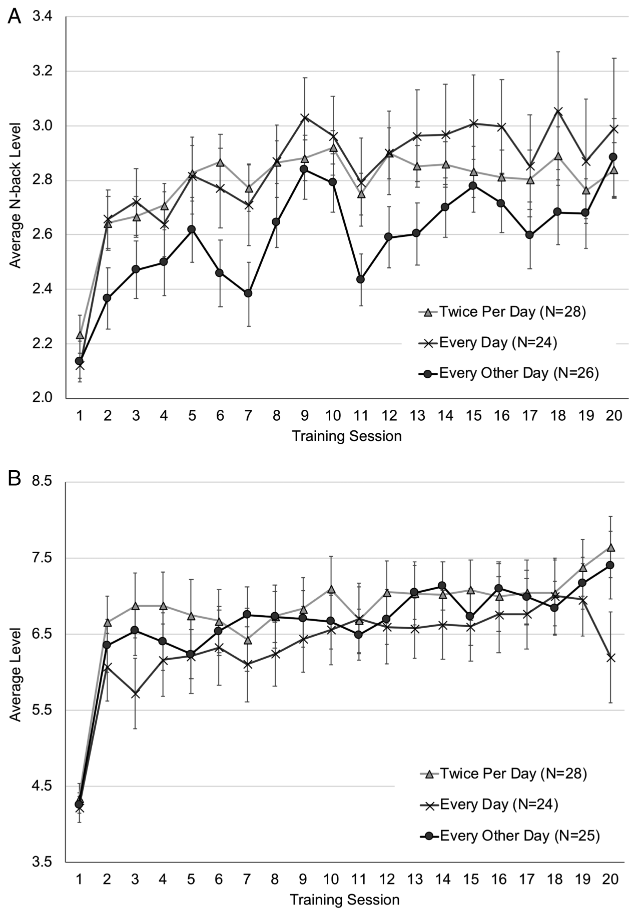

Figure 2. Training performance as a function of spacing condition and intervention group. (A) Working memory-training group. (B) Knowledge-training group. Error bars represent standard errors of the mean.

Supplementary Tables 3 and 4. There were no group differences in performance at pre-test for any of the variables (all *p*s > .11; all *BF*s < 0.18).

A comprehensive multivariate analysis of variance (MANOVA) using gain scores (pre vs post, and pre vs follow-up) that included all outcome measures (accuracy) for which we expected transfer resulted in overall significant effects of intervention in favor of the WM-training group at both, post-test, and follow-up (*pre vs post*: $F(9,125) = 3.30$, $p = .001$, $\eta^2_p = .19$; *pre vs follow-up*: $F(9,81) = 2.48$, $p = .015$, $\eta^2_p = .22$). Adding spacing group as additional variable did not change the results, and neither the main effects of spacing group, nor the Intervention × Spacing Condition interactions were significant (all *p*s > .06).

Furthermore, we calculated analyses of covariance (ANCOVA) using post-test/follow-up performance as the dependent variable, group (WM vs knowledge training) as a between-subjects factor, and pre-test performance as a covariate. Spacing condition was included as an additional between-subjects factor. The results for the various cognitive domains are reported in the main text later, and in addition, the results for each individual measure are provided in Tables 2 and 3. (In the Supplementary Material, we also report the results of MANOVAs that include the gain scores for each of the measures within a construct.)

A confirmatory factor analysis revealed that a 3-factor model represented our data adequately (Supplementary Figure 4A), and thus, we tested for intervention effects with ANCOVAs using the composites (z-scored) of the post-tests as dependent variables and pre-tests (z-scored composite) as covariates for each of the three constructs (WM, inhibition/interference, episodic memory). (Note that a 2-factor model where the WM and inhibition/interference measures were combined into one construct represented our data equally well [cf. Supplementary Figure 4B for details].) In addition, we conducted separate analyses for the WM reaction time (RT) measures (n-back and Sternberg), and control measures (processing speed, vocabulary). See Supplementary Material for more details.

For the WM accuracy composite, the WM-training group outperformed the knowledge-training group at post-test ($F(1,141) = 3.98$, $p = .02$; $\eta^2_p = .03$; $BF_{10} = 1.05$), a small effect that was reduced at follow-up ($F(1,95) = 2.29$, $p = .07$; $\eta^2_p = .02$; $BF_{10} = 0.56$). Although there was an overall effect on the composite at post-test, this was driven mostly by the n-back task, which showed strong effects in both post and follow-up sessions (cf. Tables 2 and 3). In the WM RT composite, the WM-training group similarly outperformed the knowledge-training group (post: $F(1,146) = 7.33$, $p = .004$; $\eta^2_p = .05$; $BF_{10} = 5.06$), and this effect, though small, was sustained at follow-up ($F(1,99) = 3.07$, $p = .04$; $\eta^2_p = .03$; $BF_{10} = 0.22$). The improvement here was mostly driven by the Sternberg task, which—by itself—showed particularly strong effects at the follow-up test (cf. Tables 2 and 3). The addition of spacing condition as a between-subject factor in the ANCOVA did not change any of the results, and moreover, there were no main effects of spacing or Intervention × Spacing interactions in either of the composite measures (all *p*s > .26; all BF_{10}s < 0.09).

For the inhibition/interference composite, the WM-training group outperformed the knowledge-training group at post-test, ($F(1,145) = 4.12$, $p = .02$; $\eta^2_p = .03$; $BF_{10} = 1.18$), controlling for pre-test performance, and the effect was similar at follow-up ($F(1,98) = 4.96$, $p = .01$; $\eta^2_p = .05$; $BF_{10} = 1.08$), although it was small in both cases. The improvements at post-test were mainly driven by the n-back false alarm rate, but at follow-up, participants in the WM-training group also outperformed the knowledge-training group in the D2 measure to some extent (cf. Tables 2 and 3). Again, adding spacing condition as a factor in the analyses did not change the results, and there were no effects of spacing (all *p*s > .17; all BF_{10}s < 0.13). (When using the combined WM/interference composite (accuracy; Supplementary Figure 4B), the WM-training group also outperformed the knowledge-training group at post-test controlling for pre-test performance ($F(1,137) = 6.81$, $p = .005$; $\eta^2_p = .05$; $\mathrm{BF}_{10} = 3.73$), and the effect was similar at follow-up ($F(1,93) = 4.49$, $p = .02$; $\eta^2_p = .05$; $BF_{10} = 1.43$).

Table 1. Training Data as a Function of Intervention Group and Spacing Condition

	Every other day		Every day		Twice per day		Main effect (spacing)		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	*p*	*BF*	η^2_p
Working memory-training group	*N* = 26		*N* = 24		*N* = 28				
Number of sessions	19.46	2.14	20.46	2.13	21.61	3.78			
Range of sessions	14–26		17–26		14–29				
Average first two sessions	2.25	0.45	2.39	0.35	2.44	0.42	.24	0.35	.04
Average last two sessions	2.65	0.53	2.91	1.10	2.88	0.50	.39	0.23	.03
Overall gain	0.39	0.41	0.52	0.97	0.44	0.41	.79	0.13	.01
Linear regression									
Overall slope	0.02	0.02	0.03	0.05	0.01	0.02	.37	0.24	.03
Intercept	2.39	0.47	2.57	0.48	2.64	0.39	.13	0.58	.05
Spline regression									
First slope	0.21	0.25	0.40	0.43	0.27	0.24	.09	0.73	.06
Change from first slope	−0.18	0.29	−0.41	0.39	−0.29	0.27	*	**1.20**	.08
Second slope	0.03	0.09	0.00	0.07	−0.01	0.13	.35	0.26	.03
Knot location	6.20	3.91	4.75	3.14	5.48	3.74	.38	0.24	.03
Knowledge-training group	*N* = 25		*N* = 24		*N* = 28				
Number of sessions	20.36	1.85	19.88	1.70	21.29	3.54			
Range of sessions	15–25		16–26		13–30				
Average first two sessions	5.31	1.10	5.14	1.38	5.50	1.32	.61	0.17	.01
Average last two sessions	7.07	1.83	6.74	2.25	7.18	2.19	.73	0.14	.01
Overall gain	1.77	1.27	1.59	1.10	1.68	1.25	.88	0.12	.00
Linear regression									
Overall slope	0.08	0.06	0.07	0.05	0.06	0.05	.56	0.18	.02
Intercept	5.85	1.37	5.57	1.86	6.14	1.86	.49	0.20	.02
Spline regression									
First slope	1.54	1.39	1.45	1.87	1.88	1.54	.61	0.17	.01
Change from first slope	−1.61	1.38	−1.47	1.88	−1.77	1.68	.81	0.13	.01
Second slope	−0.07	0.45	−0.02	0.26	0.10	0.29	.18	0.43	.05
Knot location	5.12	4.07	5.63	4.19	4.37	3.68	.53	0.19	.02

Note: *BF* = Bayes factors. *BF* above 1 (in bold) are considered evidence in favor of the alternative hypothesis. *P = .05.

The addition of spacing condition as a factor in the analyses did not change the results, and there were no effects of spacing [all *p*s > .27; all BF_{10}s < 0.14].)

In the episodic memory composite, there were no group differences at post-test, controlling for pre-test performance ($F(1,142) = 0.04$, $p = .42$; $\eta^2_p < .001$; $BF_{10} = 0.18$), and neither was there an effect at follow-up ($F(1,95) = 0.63$, $p = .21$; $\eta^2_p = .007$; $BF_{10} = 0.28$). Again, adding spacing condition to the analyses did not change the results, and there were no effects of spacing either (all *p*s > .23; all BF_{10}s < 0.04).

As predicted, there were no effects for the control measures favoring the WM- or knowledge-training group (*vocabulary*: post: $F(1,150) = 2.98$, $p = .09$; $\eta^2_p = .02$; $BF_{10} = 1.09$; follow-up: $F(1,99) = 1.64$, $p = .20$; $\eta^2_p = .02$; $BF_{10} = 0.57$; *processing speed*: $F(1,149) = 1.41$, $p = .24$; $\eta^2_p = .009$; $BF_{10} = 0.45$; follow-up: $F(1,104) = 1.62$, $p = .21$; $\eta^2_p = .02$; $BF_{10} = 0.40$). Adding spacing condition to the analyses did not change the results, and there were no effects of spacing either (all *p*s > .26; all BF_{10}s < 0.15).

Discussion

In this study, we tested whether the distribution of training sessions, that is, the spacing of training, plays a role in cognitive training performance and/or transfer in a healthy older adult population. Using three different training schedules (training twice per day, once every day, or once every-other-day), participants completed either an intervention targeting WM, or an alternative intervention that required fact and vocabulary learning over the course of 20 sessions. Our data provide little evidence that spacing plays any role in learning, at least with the fairly restricted spacing schedule, or the training tasks and outcome measures used here, neither when using the full sample, nor with a sample limited only to the participants who adhered strictly to the assigned spacing protocol.

Although descriptively, the participants in the every-other-day condition seem to underperform during training compared to the other two spacing conditions within the WM-training group (Figure 2A), potentially reflecting age-related deficits in consolidation (Spencer, Gouw, & Ivry,

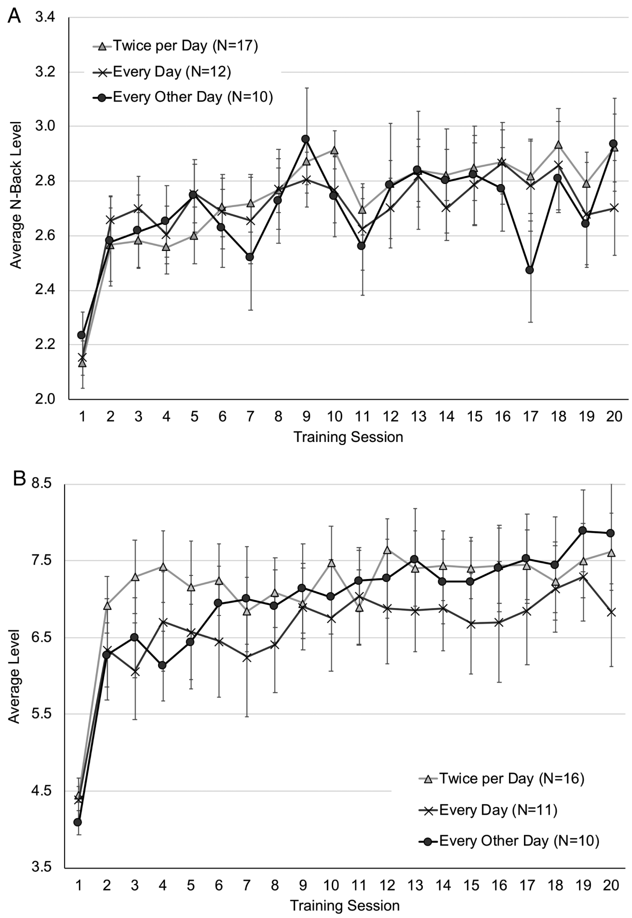

Figure 3. Training performance as a function of spacing condition and intervention group in the restricted sample that includes only participants who followed the spacing schedule as defined by using a conservative classification approach (cf. Supplementary Material). (A) Working memory-training group. (B) Knowledge-training group. Error bars represent standard errors of the mean.

2007), we did not detect any reliable differences, at least with the measures, and statistical approaches used here. Furthermore, those differences disappear in the reduced sample that strictly adhered to the spacing protocol (Figure 3a), suggesting that factors other than spacing might be at play there.

In terms of transfer, overall, we did observe effects in favor of the WM-training group immediately after training completion, as well as 3 months after training completion as evidenced by a comprehensive MANOVA that included all accuracy measures. When using theoretically derived composites, we observed that the improvements were driven by non-trained measures of WM and inhibition/interference, and thus, our data are consistent with previous work in older adults (Borella et al., 2013; Stepankova et al., 2014). We observed transfer to WM using both a composite measure representing accuracy and another that represented reaction times. Although accuracy effects were exclusively driven by the WM-training group's improvements in a non-trained spatial variant of the n-back task, the improvements in reaction time were driven by the WM-training group's performance in a Sternberg item recognition task. Importantly, the improvements in those individual measures remained present at the 3-month follow-up, indicating that the effects are not simply short lived. In addition to the observed transfer to WM functions, the WM-training group demonstrated improvements in a composite representing inhibition/interference, which were mainly driven by reduced false alarm rates in the non-trained variant of n-back, but also by a more pronounced improvement in the D2 measure, especially at follow-up. Given that our intervention deliberately incorporated the requirement to resolve interference through lure trials, the transfer effects observed in the n-back task might indicate improvements in memory precision, whereas the improvements in D2 are more indicative of changes in inhibitory control.

The transfer effects we observed in WM and inhibition/interference are consistent with our previous neuroimaging work in young adults demonstrating that WM measures that also include an interference component share overlapping neural networks (Hsu, Jaeggi, & Novick, 2017), and as such, our observed transfer effects along with the finding that the WM and inhibition/interference measures can be represented by the same latent construct provide behavioral evidence for that notion.

In contrast to previous work, however, there was no evidence of transfer to measures of episodic memory. Although we did observe improvements in a visual LTM measure successfully used in previous intervention work (Buschkuehl et al., 2008), in this study, both the WM- and knowledge-training groups improved to a similar extent. Thus, despite the relatively large effect sizes in both groups, it is not clear whether those improvements go beyond test–retest effects due to the fact that we do not have a no-contact control group. No improvements were evident in either of the other episodic memory measures, including the one that represented everyday memory functioning (CEDAR). Thus, we conclude that the n-back intervention used here might not be an ideal vehicle for improving such skills and that other approaches, such as those that include a focus on metacognitive strategies, might be more promising (e.g., Strickland-Hughes, 2017; Vranić, Španić, Carretti, & Borella, 2013).

Several limitations of this study should be noted. Specifically, due to the design of the study, the participants in the three spacing conditions completed their pre- and post-test assessments over different intervals. That is, while it took the twice-per-day group only two weeks to complete the intervention, the every-other-day group took three times as long. This difference may contribute to the lack of spacing effects in the transfer measures because memory effects due to retesting might be stronger in groups with shorter intervals between assessments, and as such, any benefits of spacing may have been overridden by the time between pre-test and post-test. However, the pre-post time interval should be less relevant at the 3-month follow-up, where spacing effects were expected to be larger, yet none were

Table 2. Intervention Effects at Post-test (Analyses of Covariances [ANCOVAs]; Pre-test Performance as Covariate)

Outcome measure	*df*	*F*	*p*	η^2_p	*BF*
Working Memory measures					
Spatial n-back (Acc)[a]	1, 150	29.97	***	.17	**63,000.43**
Sternberg (Acc)	1, 149	0.10	.76	.00	0.18
Symmetry Span (Acc)	1, 145	0.03	.88	.00	0.18
Spatial n-back (RT, correct)[a]	1, 150	0.71	.40	.01	0.12
Sternberg (RT, correct)	1, 149	8.09	**	.05	**6.57**
Inhibitory control measures					
Spatial n-back (false alarms)	1, 149	10.87	***	.07	**21.85**
D2 (*TN* – *E*)	1, 151	1.47	.23	.01	0.35
VLTM (intrusions)	1, 149	0.20	.66	.00	0.18
Long-term memory measures					
Metamemory (recall)	1, 146	0.01	.94	.00	0.18
VLTM (recall)	1, 149	0.28	.60	.00	0.20
CEDAR (Acc)	1, 151	0.63	.43	.00	0.23
Control measures					
Mill Hill Vocabulary	1, 150	2.98	.09	.02	0.70
Letter comparison	1, 149	0.45	.50	.00	0.21
Pattern comparison	1, 151	1.35	.25	.01	0.31

Notes: p values and Bayes factors (*BF*) are two tailed (uncorrected). BF above 1 (in bold) are considered evidence in favor of the alternative hypothesis. Acc = accuracy; CEDAR = Characterization of the Elderly on Daily Activities in the Real-World; RT = reaction time; VLTM = visual long-term memory.
[a]Assumptions for ANCOVA were not met, thus, the values for the Session × Group interaction (repeated measures analysis of variance) are reported instead.
$^{**}p < .01$. $^{***}p < .001$ (two tailed).

Table 3. Intervention Effects at Follow-up (Analyses of Covariances [ANCOVAs]; Pre-test Performance as Covariate)

Outcome measure	*df*	*F*	*p*	η^2_p	*BF*
Working memory measures					
Spatial n-back (Acc)	1, 101	15.51	***	.13	**144.68**
Sternberg (Acc)	1, 102	0.00	.95	.00	0.21
Symmetry span (Acc)	1, 99	0.00	.96	.00	0.21
Spatial n-back (RT, correct)	1, 101	0.04	.84	.00	0.21
Sternberg (RT, correct)	1, 102	10.54	**	.09	**19.87**
Inhibitory control measures					
Spatial n-back (false alarms)	1, 101	10.19	**	.09	**16.01**
D2 (*TN* – *E*)	1, 101	3.49	.07	.03	0.97
VLTM (intrusions)	1, 103	0.02	.88	.00	0.21
Long-term memory measures					
Metamemory (recall)	1, 97	0.00	.97	.00	0.22
VLTM (recall)	1, 103	0.08	.78	.00	0.22
CEDAR (Acc)	1, 103	0.53	.47	.01	0.26
Control measures					
Mill Hill Vocabulary[a]	1, 99	0.38	.54	.00	0.25
Letter comparison	1, 104	0.89	.35	.01	0.30
Pattern comparison	1, 104	2.49	.12	.02	0.62

Notes: p values and Bayes factors (*BF*) are two tailed (uncorrected). BF above 1 (in bold) are considered evidence in favor of the alternative hypothesis. Acc = accuracy; CEDAR = Characterization of the Elderly on Daily Activities in the Real-World; RT = reaction time; VLTM = visual long-term memory.
[a]Assumptions for ANCOVA were not met, thus, the values for the Session × Group interaction (repeated measures analysis of variance) are reported instead.
$^{**}p < .01$. $^{***}p < .001$ (two tailed).

observed for any measures. Nonetheless, a design that controls for the timing between training and assessment sessions could shed more light on this issue.

Another limitation is the restricted range of spacing conditions used here. More extreme spacing conditions (e.g., conditions that require completing all 20 sessions within 2 days, vs completing only one session per week) might have led to performance differences that better reflect massed versus distributed learning. Such a design has its own logistical challenges though, including the issue of time between assessments mentioned earlier. Alternatively, combining massed and spaced approaches within one group might be beneficial, especially in an older population. Specifically, the first few sessions could be held in a massed protocol in order to maximize task familiarization and early learning (Lövdén et al., 2010), and over time, the sessions could be increasingly spaced to gauge the actual effects of distributed learning. Such an approach has intuitive appeal; however, the relevant literature is limited, conflicting (Goedert & Miller, 2008; Hauptmann, Reinhart, Brandt, & Karni, 2005), and has not examined the types of tasks trained here.

Nonetheless, it is important to recognize that our results might only apply to the specific form of cognitive training used here (WM and knowledge training), and that the lack of measurable spacing effects is restricted to a relatively narrow range of spacing conditions. It is critical for future research to test different variations of training schedules using other interventions (e.g., speed of processing training) over longer periods of time with appropriate outcome measures, and to replicate the present results with populations that differ in demographic and health characteristics (e.g., young adults).

Nonetheless, our results have practical implications. Even though our older participants were asked to perform a demanding cognitive intervention unsupervised and independently at home, they were not only able to complete their intervention and showed considerable task-specific improvements, but also, they were reasonably compliant. Almost everyone who started the intervention completed the requested number of training sessions. In fact, many participants (especially those in the twice-per-day condition) trained more than the required amount, some training more than three times the required sessions, resulting in their exclusion from the analyses (cf. Figure 1). Our data reflect what participants might be doing in the real world, for example, when they choose to complete cognitive training on their own. It seems that even complex interventions do not necessarily have to be conducted in the lab following a precise training schedule in order to be effective.

Conclusions and Implications

Our study demonstrated that healthy older adults are able to complete a fairly complex intervention on tablet devices independently at home with minimal researcher interaction (see also Stepankova et al., 2014). Furthermore, training WM seems to have some benefits in terms of improving performance in non-trained measures, in particular, WM and inhibition/interference, which are sustained up to 3 months after training completion. However, in contrast to other studies, our intervention did not result in improvements in episodic memory including a measure that reflects real-world requirements.

Most critically, our results indicate that the spacing of training sessions seems to have negligible effects on training outcome, at least within the restricted range of spacing conditions used here, and that other factors may be more important for producing more robust training and transfer outcomes.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

This work was supported by the National Institute on Aging (1R01AG049006 to S. M. Jaeggi, J. Jonides, P. Reuter-Lorenz, and P. Shah, and 1K02AG054665 to S.M. Jaeggi). C. M. Parlett-Pelleriti is supported by the National Science Foundation, Division of Graduate Education (Grant No. 1849569).

Conflict of Interest

M. Buschkuehl is employed at the MIND Research Institute whose interest is related to this work. S. M. Jaeggi has an indirect financial interest in the MIND Research Institute.

References

Alzheimer's Association. (2015). 2015 Alzheimer's disease facts and figures. *Alzheimer's and Dementia,* **11**, 332–384. doi:10.1016/j.jalz.2015.02.003

Borella, E., Carretti, B., Zanoni, G., Zavagnin, M., & De Beni, R. (2013). Working memory training in old age: An examination of transfer and maintenance effects. *Archives of Clinical Neuropsychology,* **28**, 331–347. doi:10.1093/arclin/act020

Brehmer, Y., Rieckmann, A., Bellander, M., Westerberg, H., Fischer, H., & Bäckman, L. (2011). Neural correlates of training-related working-memory gains in old age. *Neuroimage,* **58**, 1110–1120. doi:10.1016/j.neuroimage.2011.06.079

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience,* **6**, 63. doi:10.3389/fnhum.2012.00063

Brickenkamp, R. (2002). *Test d2. The d2 test of attention* (9th ed.). Goettingen, Germany: Hogrefe.

Buckner, R. L. (2004). Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. *Neuron,* **44**, 195–208. doi:10.1016/j.neuron.2004.09.006

Bugg, J. M., Zook, N. A., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain and Cognition,* **62**, 9–16. doi:10.1016/j.bandc.2006.02.006

Buschkuehl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Dapp, C., Muller, M., . . . Perrig, W. J. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology and Aging,* **23**, 743–753. doi:2008-19072-008

Cahn-Weiner, D. A., Farias, S. T., Julian, L., Harvey, D. J., Kramer, J. H., Reed, B. R., . . . Chui, H. (2007). Cognitive and neuroimaging predictors of instrumental activities of daily living. *Journal of the International Neuropsychological Society,* **13**, 747–757. doi:10.1017/S1355617707070853

Cantarella, A., Borella, E., Carretti, B., Kliegel, M., & de Beni, R. (2017). Benefits in tasks related to everyday life competences after a working memory training in older adults. *International Journal of Geriatric Psychiatry,* **32**, 86–93. doi:10.1002/gps.4448

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

Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science (New York, N.Y.),* **320**, 1510–1512. doi:10.1126/science.1155466

Diehl, M., Marsiske, M., Horgas, A. L., Rosenberg, A., Saczynski, J. S., & Willis, S. L. (2005). The revised observed tasks of daily living: A performance-based assessment of everyday problem solving in older adults. *Journal of Applied Gerontology,* **24**, 211–230. doi:10.1177/0733464804273772

Fang, J., Power, M., Lin, Y., Zhang, J., Hao, Y., & Chatterji, S. (2012). Development of short versions for the WHOQOL-OLD module. *Gerontologist,* **52**, 66–78. doi:10.1093/geront/gnr085

Flegal, K. E., Ragland, J. D., & Ranganath, C. (2019). Adaptive task difficulty influences neural plasticity and transfer of training. *NeuroImage,* **188**, 111–121. doi:10.1016/j.neuroimage.2018.12.003

Flegal, K. E., & Reuter-Lorenz, P. A. (2014). Get the gist? The effects of processing depth on false recognition in short-term and long-term memory. *Memory and Cognition,* **42**, 701–711. doi:10.3758/s13421-013-0391-9

Goedert, K. M., & Miller, J. (2008). Spacing practice sessions across days earlier rather than later in training improves performance of a visuomotor skill. *Experimental Brain Research,* **189**, 189–197. doi:10.1007/s00221-008-1414-9

Hauptmann, B., Reinhart, E., Brandt, S. A., & Karni, A. (2005). The predictive value of the leveling off of within session performance for procedural memory consolidation. *Brain Research. Cognitive Brain Research,* **24**, 181–189. doi:10.1016/j.cogbrainres.2005.01.012

Hill, N. T., Mowszowski, L., Naismith, S. L., Chadwick, V. L., Valenzuela, M., & Lampit, A. (2017). Computerized cognitive training in older adults with mild cognitive impairment or dementia: A systematic review and meta-analysis. *American Journal of Psychiatry,* **174**, 329–340. doi:10.1176/appi.ajp.2016.16030360

Hsu, N. S., Jaeggi, S. M., & Novick, J. M. (2017). A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language,* **166**, 63–77. doi:10.1016/j.bandl.2016.12.006

Iordan, A. D., Cooke, K. A., Moored, K. D., Katz, B., Buschkuehl, M., Jaeggi, S. M., . . . Reuter-Lorenz, P. A. (2018). Aging and network properties: Stability over time and links with learning during working memory training. *Frontiers in Aging Neuroscience,* **9**, 419. doi:10.3389/fnagi.2017.00419

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory and Cognition,* **42**, 464–480. doi:10.3758/s13421-013-0364-z

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive-control and working memory training in older adults. *Psychological Science,* **25**, 2027–2037. doi:10.1177/0956797614548725

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., & Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Frontiers in Human Neuroscience,* **8**, 242. doi:10.3389/fnhum.2014.00242

Krug, D., Davis, T. B., & Glover, J. A. (1990). Massed versus distributed repeated reading—A case of forgetting helping recall. *Journal of Educational Psychology,* **82**, 366–371. doi:10.1037/0022-0663.82.2.366

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin,* **136**, 659–676. doi:10.1037/a0020080

Lustig, C., Shah, P., Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychology Review,* **19**, 504–522. doi:10.1007/s11065-009-9119-9

McGillivray, S., & Castel, A. D. (2011). Betting on memory leads to metacognitive improvement by younger and older adults. *Psychology and Aging,* **26**, 137–142. doi:10.1037/a0022681

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

Parlett-Pelleriti, C., Lin, G. C., Jones, M. R., Linstead, E., & Jaeggi, S. M. (2019). Exploring age-related metamemory differences using modified brier scores and hierarchical clustering. *Open Psychology,* **1**, 215–238. doi:10.1515/psych-2018-0015

Perrig, W. J., Etienne, A., Jaeggi, S. M., Blaser, D., Meier, B., Hofer, D., . . . Buschkuehl, M. (2006). *Computerunterstützter Gedächtnis-Funktion-Test (C-GFT) (Version 3.1)*. Bern: Universität Bern.

Raven, J., Raven, J., & Court, R. (1998). *Mill hill vocabulary scale*. Oxford: Oxford University Press.

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H-Y., King, J. W., . . . Willis, S. L. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatrics Society,* **62**, 16–24. doi:10.1111/jgs.12607

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment,* **28**, 164–171. doi:10.1027/15-5759/A000123

Ribaupierre, A. de, & Lecerf, T. (2006). Relationships between working memory and intelligence from a developmental perspective: Convergent evidence from a neo-Piagetian and a psychometric approach. *European Journal of Cognitive Psychology,* **18**, 109–137. doi:10.1080/09541440500216127

Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and Aging,* **26**, 813–822. doi:10.1037/a0023631

Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanism of active maintenance and executive control* (pp. 1–26). New York: Cambridge University Press.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest,* **17**, 103–186. doi:10.1177/1529100616661983

Spencer, R. M., Gouw, A. M., & Ivry, R. B. (2007). Age-related decline of sleep-dependent consolidation. *Learning and Memory (Cold Spring Harbor, N.Y.),* **14**, 480–484. doi:10.1101/lm.569407

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine,* **166**, 1092–1097. doi:10.1001/archinte.166.10.1092

Stepankova, H., Lukavsky, J., Buschkuehl, M., Kopecek, M., Ripova, D., & Jaeggi, S. M. (2014). The malleability of working memory and visuospatial skills: A randomized controlled study in older adults. *Developmental Psychology,* **50**, 1049–1059. doi:10.1037/a0034913

Strickland-Hughes, C. M. (2017). *Memory training in adulthood: Changing memory and self-regulation* (PhD, University of Florida). Retrieved on December 24, 2018, from https://search.proquest.com/docview/2068015357/abstract/8D9B1CB97FA348E1PQ/1.

Thomas, K. R. (2015). *Understanding errors in complex everyday cognitive tasks in older adults*. (PhD) (Doctoral dissertation, University of Florida).

Tomaszewski Farias, S., Cahn-Weiner, D. A., Harvey, D. J., Reed, B. R., Mungas, D., Kramer, J. H., & Chui, H. (2009). Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. *Clinical neuropsychologist,* **23**, 446–461. doi:10.1080/13854040802360558

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society,* **40**, 922–935. doi:10.1111/j.1532-5415.1992.tb01992.x

Vranić, A., Španić, A. M., Carretti, B., & Borella, E. (2013). The efficacy of a multifactorial memory training in older adults living in residential care settings. *International Psychogeriatrics,* **25**, 1885–1897. doi:10.1017/S1041610213001154

Wang, Z., Zhou, R., & Shah, P. (2014). Spaced cognitive training promotes training transfer. *Frontiers in Human Neuroscience,* **8**, 217. doi:10.3389/fnhum.2014.00217

Weicker, J., Villringer, A., & Thöne-Otto, A. (2016). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology,* **30**, 190–212. doi:10.1037/neu0000227

Yesavage, J. A. (1988). Geriatric depression scale. *Psychopharmacology Bulletin,* **24**, 709–711. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/3249773. PMID: 3249773.