Journal of Cognitive Enhancement
https://doi.org/10.1007/s41465-025-00342-4

RESEARCH

Unlocking Cognitive Potential: How Working Memory Training Impacts Reading Skills in Aging Adults

Sara Assecondi[1]

Received: 23 March 2025 / Accepted: 10 December 2025

Abstract
Working memory is crucial for maintaining independence in daily activities, especially as we age. Cognitive programs aim to enhance cognitive performance to support independent living, but the transfer of improvements from these exercises to daily activities remains unclear. This study uses reading comprehension, a complex activity involving information storage and processing, as a proxy for everyday functions. I examine the relationship between visuo-spatial working memory and reading comprehension, and whether training-related improvements in working memory capacity translate to reading performance. In a sample of 175 individuals undergoing a 5-day cognitive training program, I found that the extent to which improvements in working memory transfer to reading comprehension may differ descriptively across age groups, with preliminary evidence of a modest relationship in younger adults but not in older adults. Thus, standard training programs may not lead to noticeable improvements in real-life tasks, indicating the need for more ecologically valid measures. This knowledge can help in designing better and more effective training programs to counteract cognitive decline.

Keywords Working memory · Cognitive training · Reading comprehension · Aging · Transfer effects

Introduction

Physiological changes in cognitive functioning come with aging. For some, these changes are minor nuisances, but for others, they can significantly impact independence and quality of life (Cohen et al., 2019). The rise of validated cognitive training programs offers a potential solution to improve cognition (Borella et al., 2017; Hou et al., 2020; Wang et al., 2021). However, their effectiveness, especially for healthy aging, remains debated, with several meta-analyses and reviews reporting mixed or null transfer of benefits to untrained functions (Gates et al., 2019; Gobet & Sala, 2023; Melby-Lervåg & Hulme, 2013; Sala et al., 2019; Simons et al., 2016).

There is a discrepancy between what individuals expect from cognitive training programs and what these programs achieve (Ng et al., 2020). Cognitive training programs are typically designed using standardized laboratory tests to assess specific cognitive functions and are then adapted, sometimes with gamified elements to encourage participation, for better compliance (Vermeir et al., 2020). However, participants often seek out these programs due to self-reported memory complaints related to everyday activities. This creates a mismatch between the tools used for assessment and the cognitive functions of interest (Goghari & Lawlor-Savage, 2018; Osborne-Crowley, 2020; Shamay-Tsoory & Mendelsohn, 2019; Sunderland et al., 1983).

Cognitive programs often target core functions, such as working memory, to enhance both the trained function and related activities (Katz et al., 2018). The rationale is that core cognitive functions support higher-level tasks and daily living, thus, improvements in e.g., working memory would lead to overall better performance in daily activities.

One such daily activity is reading comprehension, a critical skill essential for independent living and the execution of everyday tasks (Gordon et al., 2016). This study investigates the relationship between working memory and reading comprehension and whether common working memory training programs can improve reading performance in a naturalistic task.

✉ Sara Assecondi
Sara.assecondi@gmail.com; sara.assecondi@unint.it

[1] Center for Mind/Brain Sciences - CIMeC, University of Trento, Rovereto, TN, Italy

Published online: 04 January 2026

Working Memory Training in Older Adults

Cognitive training programs frequently target working memory, a fundamental cognitive function, typically declining with age, that is responsible for temporarily storing and manipulating information (Naveh-Benjamin & Cowan, 2023). Working memory training programs have been successful in enhancing working memory performance in older adults (Borella et al., 2017; Simon et al., 2018). However, transfer effects are rarely achieved (Borella et al., 2017; Melby-Lervåg et al., 2016; Richmond et al., 2011; Sala et al., 2019; Waris et al., 2015; Zelinski, 2009), and improvements in daily activities are seldom quantified (Fan & Wong, 2019; Rebok et al., 2014; Richmond et al., 2011; Tennstedt & Unverzagt, 2013; Willis et al., 2006).

Working memory training programs target two key components of working memory: *storage* and *updating* (or *manipulation*). In lab settings and animal studies, the *storage* component is often measured using a visuo-spatial delayed match-to-sample task (DMTS, Vogel & Machizawa, 2004), and DMTS training has shown task-specific improvements in young (Buschkuehl et al., 2017) and older adults (Tagliabue et al., 2024). The *updating* component of working memory is often measured with tasks like the n-back (NB), engaging the visual, verbal domains, or both (Miller et al., 2009; Verhaeghen & Basak, 2005), depending on the materials used.

Reading Comprehension and Working Memory

There are divergent views regarding the relationship between working memory and reading comprehension. Some studies suggest that the observed association may be largely due to collinearity with other cognitive abilities (Caplan & Waters, 2013; Van Dyke et al., 2014). In contrast, other research emphasizes that working memory, particularly its storage and updating capacities, is fundamentally linked to reading comprehension (Daneman & Carpenter, 1980; DeCaro et al., 2016; Just & Carpenter, 1992; Kim et al., 2016; Norman et al., 1992).

Martin et al. (Martin et al., 2020) proposed that reading comprehension performance relies on two aspects of information processing, not only the ability to maintain (*storage*) relevant information but also in the role of forgetting (or *updating* information when new information becomes available). This is in agreement with research showing that complex span tasks are better predictors of reading performance than simple span (Daneman & Carpenter, 1980; Turner & Engle, 1989), because of the engagement of more complex processing in addition to storage (Daneman & Merikle, 1996; Martin et al., 2020).

Given the link between reading comprehension and working memory and the fact that working memory declines with age, one would expect a similar trend for reading comprehension performance. This relationship is, however, not that simple (Radvansky et al., 2001). The link between WM and reading comprehension may depend on individual cultural background and reading efficiency, as well as meta-comprehension abilities (De Beni et al., 2007; Martín-Aragoneses et al., 2023). In addition, older adults often adopt strategies to mitigate working memory deficits.

The relation between working memory and reading comprehension might also change with age, depending on the type of reading demands, showing indeed that despite a significant decline in working memory performance, differences in reading skills are not that consistent between young and older adults (De Beni et al., 2003). Peng et al. (2018) highlighted that inexperienced younger children with less decoding and semantic retrieval capabilities may rely more on WM than more experienced readers. This relation may remain true between young and older adults, where experience-rich backgrounds in older adults may reduce their reliance on WM for reading compared to younger adults.

When reading a text, individuals should master two mechanisms: *parsing*, that is, identifying units of meaning within a sentence, and *interpretation*, that is, understanding the meaning of a sentence (Caplan & Waters, 2013), mechanisms that rely on storing, retrieving, and updating information in memory, thus sensitive to working memory capacity. It is, therefore, not so far-fetched to think that if training can be used to improve working memory, reading comprehension might also benefit from these changes.

Considering the variety of working memory training programs, to understand the potential of training to improve reading skills, one should consider the "*domain-specificity debate of working memory*" (Peng et al., 2018), that is, whether training in different working memory domains (verbal or visual) has a differential impact on reading performance. The question stems from the contrasting view of working memory as a *domain-specific* or *domain-general model*. Many researchers supporting the "*domain-general"* view propose that the central executive, as the core component of working memory, coordinates the processing of visual, visuo-spatial, and verbal information. Individual differences in central executive capacity have been linked to performance across these domains (Engle, 2002; Engle & Kane, 2003). Consequently, training that engages these processes, regardless of the specific material, may influence central executive functioning, suggesting that visuo-spatial working memory training could also affect reading skills. In contrast, researchers that support the *"domain-specific"* model affirm that working memory performance depends on domain-specific knowledge (Ericsson & Kintsch, 1995).

There is evidence supporting both models: in children, it has been shown that both visual and verbal working memory can predict successful reading fluency and comprehension (Peng et al., 2018), suggesting that both types of cognitive training can impact reading comprehension due to the interconnected nature of the visual and verbal components of working memory. Pham and Hasson (Pham & Hasson, 2014) reported that both verbal and visuospatial working memory significantly predicted reading skills, albeit with a stronger effect for verbal working memory, while the opposite was also found (Shah & Miyake, 1996). In the context of training, (Borella et al., 2017; Carretti et al., 2014; Payne & Stine-Morrow, 2017) it was found that cognitive training targeting verbal working memory improved trained but also untrained verbal working memory tasks and some language tasks.

The possibility of improving reading comprehension through working memory training is enticing as reading is vital for maintaining independence in everyday activities (Gordon et al., 2016). Selecting the appropriate training program is crucial for its effectiveness. Given the nature of reading comprehension, one might assume that only a verbal training program would effectively enhance reading skills.

While numerous studies have emphasized the role of verbal working memory in supporting reading comprehension, there are both theoretical and empirical grounds for examining visuo-spatial working memory as well. First, individual differences in reading comprehension are influenced not only by task modality but also by the degree of attentional control required (Carretti et al., 2009). Although verbal tasks often show larger effect sizes, visuo-spatial tasks tend to yield more consistent results across studies, as reflected in their lower heterogeneity indices (I^2). This suggests that visuo-spatial working memory, even if less predictive in magnitude, contributes to reading comprehension in a more stable manner across populations, possibly due to smaller sensitivity to individual differences in strategy use, or language proficiency. From a theoretical perspective, reading comprehension is not exclusively verbal but requires domain-general executive processes, such as attentional control, updating, and integration across sentences and contexts (Gordon et al., 2016). These processes may be equally well engaged by visuo-spatial tasks that challenge storage and updating functions.

Second, visuo-spatial training programs may offer practical advantages. Visual stimuli can be more readily adapted into interactive, game-like formats (e.g., puzzles, pattern matching) and are less dependent on language, making them easier to implement across participants with different language backgrounds. In contrast, verbal materials rely heavily on text-based exercises and may require language-specific adaptation, limiting their gamification potential. Maintaining participant motivation is critical in multi-session training studies (Redlinger et al., 2022). Gamification, while not directly improving cognitive performance, has been shown to support persistence in cognitive tasks (Deveau et al., 2015; Jayalath & Esichaikul, 2022), which could be particularly helpful for older adults at higher risk of dropout (Koivisto & Malik, 2020). That said, older adults may also be intrinsically motivated to participate in cognitive training studies, and this intrinsic motivation may partially compensate for potential engagement differences.

Taken together, this combination of theoretical considerations (the domain-general nature of working memory, the consistency of visuo-spatial effects across studies) and practical advantages (higher engagement and lower attrition) provides a rationale for selecting visuo-spatial working memory training in the present study. While the engaging format of such tasks is a practical strength for online, multi-session interventions, it is important to note that engagement alone does not guarantee transfer to untrained abilities. Prior work has repeatedly shown that, although training can produce task-specific improvements, evidence for transfer to broader skills like reading comprehension remains inconsistent (Borella et al., 2017; Melby-Lervåg et al., 2016; Richmond et al., 2011; Sala et al., 2019; Waris et al., 2015; Zelinski, 2009). Given the mixed evidence in the broader literature, it remains an open empirical question whether such training can transfer to improvements in reading comprehension when assessed with naturalistic tasks. Addressing this uncertainty represents the novel contribution of the present work.

Purpose of the Present Study

The aim of this study was to investigate whether improvements in working memory (WM) capacity through targeted training transfer to reading comprehension performance in young and older adults. To this end, I implemented two types of visuo-spatial WM training: one focusing on **storage and maintenance** (Delay Match-to-Sample, DMTS) and one focusing on **updating** (n-back, (S. M. Jaeggi et al. 2010a, b)). Participants were randomly assigned to either one of the two training conditions or to an active control group.

In the present study, the n-back task was implemented as a spatial working memory task, as participants were required to remember the positions of the stimuli on a grid. In contrast, the DMTS task was a visuo-spatial working memory task, as it required participants to remember both the colour and position of the stimuli. I selected the visuo-spatial training format because of previous encouraging results (Assecondi et al., 2021; Assecondi, Hu, Assecondi et al., 2022a, b; Assecondi, Villa-Sánchez et al., 2022) and for its more engaging game-like design, which is particularly suitable for online, multisession experiments where participants have no direct contact with the researcher. Reading skills were tested with a naturalistic task (De Beni et al., 2003; Martin et al., 2020), where participants were asked to read a series of paragraphs and answer some questions. The questions were fact-based or inferential responses (see Methods section) to assess reading comprehension in more ecologically valid conditions, capturing the multiple cognitive processes (maintenance, integration, and inference) that occur during everyday reading. To gain an understanding of how much memory failures impact the individual's everyday life, self-reported memory failures were evaluated using the Everyday Memory Questionnaire (EMQ) (Ossher et al., 2013). The EMQ was included as a self-report measure of perceived memory failures in daily contexts. Although not a direct measure of reading comprehension, it was considered relevant as a subjective index of memory functioning that could plausibly relate to individual differences in cognitive performance. However, given its self-report nature, any associations with objective measures were expected to be modest.

Our objectives were threefold: (a) to examine age-related differences in the relationship between WM capacity and reading comprehension at baseline; (b) to test whether WM training improves WM performance in a task-congruent manner (storage gains from DMTS, updating gains from n-back); (c) to assess transfer effects of WM training to reading comprehension, using a naturalistic paragraph-based comprehension task that included both fact-based and inferential questions.

I hypothesized that: Individual differences in reading comprehension performance would be partly explained by WM capacity, with age modulating this relationship. Training would lead to improvements in WM capacity, specific to the type of training (storage vs. updating). Increases in WM capacity would transfer to improvements in reading comprehension performance.

Materials and Methods

This study employed a randomized design with three groups: two training interventions and one active control. Participants were assigned either to (1) a visuo-spatial working memory **storage training** program based on the Delay Match-to-Sample (DMTS) task, (2) a **spatial updating training** program based on the n-back (NB) task, or (3) a control condition. The interventions were delivered fully online across five sessions. Reading comprehension, WM capacity (storage and updating), were assessed at baseline and post-training, while self-report questionnaires assessing everyday memory, affect, sleep quality, and quality of life were collected either one day after each training session for the intervention groups or during the corresponding online contact for the control group (see "Procedure" for details).

Participants

A convenience sample was collected online in 2021 and partly already reported (Tagliabue et al., 2022). Unlike Tagliabue et al. (2022), which focused on online and offline learning in a visuo-spatial task (DMTS), the present study examines both storage and updating training and their transfer to reading comprehension across age groups. Young (20–30 years) and older individuals (65–75 years) with normal or corrected vision, no cognitive impairments, dementia, or ongoing mental illness, and fluent in English were recruited via Prolific (www.prolific.co). Participants confirmed not using medication for mood disorders, recent alcohol, caffeine, or drug consumption, and adequate sleep (>6 h). Older adults completed the Self-Administered Tasks Uncovering Risk of Neurodegeneration (SATURN) (Bissig et al., 2020), excluding those with scores below 25. Of the 201 participants who completed the training, 40 (15 older, 25 young) were excluded for incomplete tasks, poor accuracy, or having completed the training outside the allocated time, leaving a final sample of 83 young (50 females, mean age 24.2 ± 3.2) and 92 older adults (47 females, mean age 68.0 ± 2.9). (see Fig. 1 for recruitment, allocation, and exclusion details). Descriptive statistics are in Table 1, with further details on nationality and levels of education, in the Supplementary Information Table S1. The study was approved by the University of Trento Research Ethics Committee (Protocol No.2021-041). Participants gave informed consent via Psytoolkit (https://www.psytoolkit.org, v3.4.4, Stoet, 2010, 2017) and were compensated £7.00 GBP/hour.

Procedure

The experiment took place online, with participants completing the training sessions independently over five days. In addition, they completed a baseline assessment (pre-training), an outcome assessment (post-training), and a follow-up session, but the latter was not included in this analysis. The order of tasks was randomized across baseline and post-training sessions, and each participant completed a uniquely assembled set of reading comprehension questions to avoid repetition and to balance difficulty, ensuring that no question was presented more than once (see Supplementary

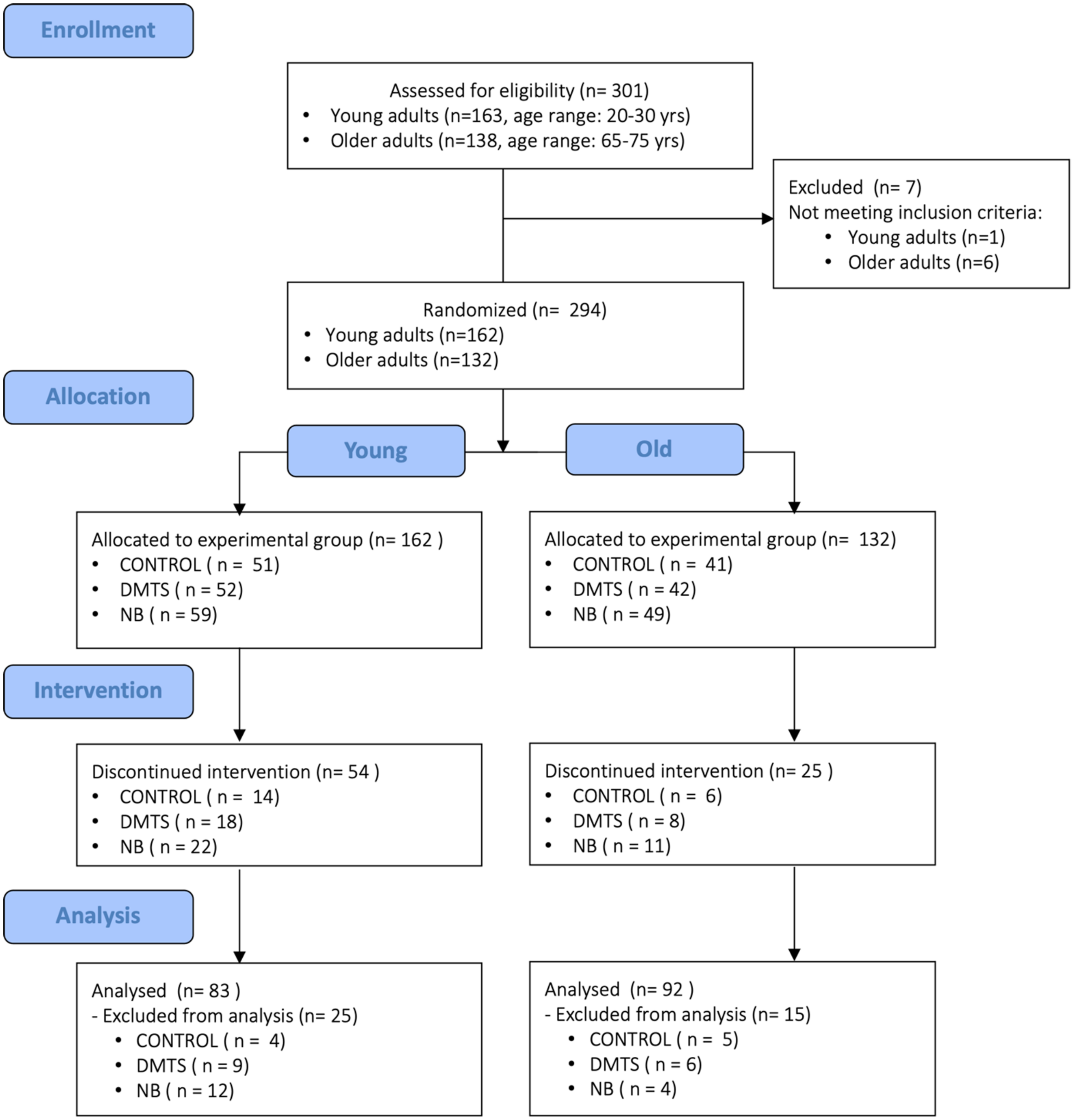

Fig. 1 Consort-like flowchart of recruitment, exclusion, and allocation of participants

Materials for details). Participants could practice beforehand and had a chance to take breaks between tasks and blocks. A diagram of the procedure is shown in Fig. 2.

Baseline assessment (pre-training). Two days before the training started, participants completed two working memory tasks (a DMTS and an n-back) to measure working memory capacity, and a naturalistic reading comprehension (RC) task. The baseline session lasted for about one hour.

Training regime. About 48 h after the baseline assessment, participants were randomized into three groups for training: a control and two working memory training groups. The control group (CONTROL) filled out one or two questionnaires online every day, following the same schedule as the training groups. One training group (DMTS) completed an adaptive delayed match-to-sample task, while the other training group (NB) completed an adaptive n-back

Table 1 Descriptive statistics for the two age groups: for each variable (except for gender), mean, standard error, and 95% confidence intervals are reported. Kst@PRE = working memory storage capacity (DMTS task) at baseline; Kup@PRE = working memory updating capacity (NB task) at baseline; %Corr@PRE = percentage of correct responses in the reading comprehension (RC) task

	Young (*n* = 83)				Old (*n* = 92)			
	M	SE	95% CI		M	SE	95% CI	
			LL	UL			LL	UL
age	24.20	0.35	23.50	24.90	68.00	0.30	67.40	68.60
gender (f/m)	50/33	--	--	--	47/45	--	--	--
education (yrs)	16.10	0.26	15.60	16.60	15.30	0.36	14.60	16.00
Saturn	--	--	--	--	27.80	0.10	27.60	28.00
Familiarity with technology	19.50	0.34	18.80	20.20	19.30	0.37	18.60	20.10
Quality of life	14.40	0.22	14.00	14.90	15.80	0.24	15.30	16.30
Positive affect	31.10	0.89	29.30	32.80	33.80	0.76	32.30	35.30
Negative affect	20.70	0.90	18.90	22.40	13.60	0.53	12.50	14.60
Pittsburg Sleep Quality Index	5.42	0.26	4.90	5.95	5.20	0.30	4.60	5.79
Everyday Memory Questionnaire	51.40	3.66	44.10	58.70	30.30	0.30	67.40	68.60
Kst@PRE	2.39	0.08	2.23	2.55	1.80	0.04	1.71	1.89
Kup@PRE	1.21	0.07	1.08	1.35	1.04	0.04	0.97	1.11
%Corr@PRE	0.87	0.01	0.86	0.89	0.88	0.01	0.87	0.89

Fig. 2 A diagram of the procedure

task before filling out the same questionnaires as the CONTROL group. Details of the tasks are reported below (section "Cognitive tasks and questionnaires" and Fig. 2). Each training session lasted for about 25 min for the DMTS and NB group, and about 5 min for the CONTROL group.

Outcome assessment. After an additional 48-hour break, participants completed the post-training assessment, with tasks presented in a different order and using items that differed from those in the baseline session. The outcome session lasted for about one hour.

Cognitive Tasks and Questionnaires

Cognitive tasks and training were programmed with PsychoPy (Peirce et al., 2019), translated into Javascript, and uploaded to Pavlovia (https://pavlovia.org/). Questionnaires were administered through Psytoolkit (https://www.psytoolkit.org/, Stoet, 2010, 2017). Participants were instructed to sit at arm's length from the screen (approximately 70 cm) and guided via on-screen instructions on how to measure this distance by extending their hand to touch the screen. Small variations in actual distance were considered unlikely to meaningfully affect task performance. Before the DMTS, participants completed a screen dimension estimation to ensure consistent stimulus size (Card Task procedure; see, (Li et al., 2020) and https://gitlab.pavlovia.org/Wake/screenscale).

Tasks are briefly described hereafter, while technical details on their implementation are reported in the **Supplementary Information**. Participants had a chance to practice the tasks, which lasted for about 8 to 10 min each.

Delayed match-to-sample task. The *delayed match-to-sample task* (DMTS) (Vogel & Machizawa, 2004) was used to measure and train working memory storage capacity (K_{st}). Colored dots were presented on a dark grey background, with targets and distractors differentiated by color. To match task difficulty across age groups (Sander et al., 2012), younger adults saw 3 to 5 colored dots per side, while older adults saw 2 to 4 dots, with 24 items total (12 per hemifield). The number of stimuli remained constant to balance sensory input. The test array followed the memory array, and individuals had to respond with the right arrow key if the test array matched the memory array, with the left arrow key otherwise (no change – change with 50–50 chance). The percentage of correct responses (%Corr) was recorded. Working memory storage capacity (K_{st}) was calculated for each set size, as $K_{st} = setsize^{*}(\mathrm{H}-\mathrm{FA})/(1-\mathrm{FA})$, where H, and FA are the fraction of hits (correct match) and false alarms (incorrect match) to match and no-match trials, respectively (Pashler, 1988), then averaged.

N-back task. The *n-back task (NB)* required the storage, update, and manipulation of information in working memory and was used to measure the spatial working memory process/updating capacity (K_{up}) (Kirchner, 1958). Stimuli consisted of a square moving on a 3×3 grid (the central square was reserved for fixation) on a dark grey background. The color of the moving square was randomized across blocks and sessions, from the pool of colors used in the DMTS task. Participants remembered the position of 'n' previous stimuli and responded by pressing the left or right arrow respectively, for 'match' (same position, 30%) or 'no match' (different position, 70%). To ensure task difficulty was comparable across age groups (Sander et al., 2012), memory loads were assigned based on previous studies (Assecondi et al., 2021; Assecondi, Hu, Assecondi et al., 2022a, b). Younger adults had 'n' values of 2, 3, or 4, while older adults had 'n' values of 1, 2, or 3. The accuracy of responses was recorded. Working memory updating capacity (K_{up}) was calculated for each 'n' as $K_{up} = n^{*}(\mathrm{H}-\mathrm{FA})$, where H, and FA are the fraction of hits (correct match) and false alarms (incorrect match) to match and no-match trials, respectively (Cowan, 2001), then averaged.

Reading comprehension. A naturalistic *reading comprehension task (RC)* was designed based on (Daneman & Carpenter, 1980) and (Martin et al., 2020). By "naturalistic assessment," I refer to a task in which participants read continuous, paragraph-length texts (~140 words) that resemble real-world reading material, such as articles or informative passages, rather than isolated sentences or artificially constrained stimuli. This approach captures reading comprehension under more ecologically valid conditions, reflecting the multiple cognitive processes (maintenance, integration, and inference) that occur during everyday reading.

Participants answered four questions per paragraph without any time limit: two *fact-based questions* (multiple choice), assessing memory for information explicitly mentioned in the paragraph, and two *inferential questions* (TRUE/FALSE/I DON'T KNOW), requiring manipulation and integration of the information presented. The use of different formats reflects the distinct cognitive demands of the two question types: fact-based questions primarily involve maintenance of information, whereas inferential questions additionally engage executive processes such as updating and manipulation, consistent with prior evidence that tasks involving storage and processing better predict individual differences in reading comprehension than storage-only tasks (Carretti et al., 2009; Daneman & Merikle, 1996). Paragraphs and questions, proofread by a native English speaker, were presented in random order; once the questions appeared, the paragraph was no longer visible. Each session included 10 paragraphs (for a total of 20 fact-based and 20 inferential questions), with no feedback provided. Paragraph difficulty was balanced across participants and sessions (see **Supplementary Information** for details). The primary dependent variable was the percentage of correct responses for each question type (%Corr). Reading time for each paragraph and reading speed (ratio of reading time to paragraph length in words) were also recorded for both conditions. This task design allowed us to examine whether training-related improvements in working memory, specifically in storage and updating capacities, would transfer to reading comprehension performance under conditions varying in cognitive demands.

Everyday Memory Questionnaire. Self-perceived memory failures in everyday life were measured with the Everyday Memory Questionnaire (EMQ) (Sunderland et al., 1984), regarding the preceding two weeks. The EMQ is a self-report instrument designed to assess the frequency of everyday memory lapses across various contexts, capturing deficits that may reflect underlying working memory impairments. It comprises items that evaluate common memory failures, with higher total scores indicating more frequent difficulties in daily memory functioning. As it targets everyday activities, it may plausibly relate to reading comprehension performance. The EMQ, which has been validated in healthy adult populations (Rodrigues et al., 2025; Royle & Lincoln, 2009; Taleb et al., 2024), supporting its use in both younger and older adults in the present study.

Other questionnaires. In addition to the EMQ, the following questionnaires were administered to each participant to track their lifestyle and habits and to assess baseline differences. *Familiarity with technology* (in-house developed), is a short in-home developed questionnaire to measure how familiar participants are with digital devices for everyday life: this is important to monitor in online studies when data quality relies on participants using a computer, especially with older populations. The *Pittsburgh Sleep Quality Index* (PSQI, Buysse et al., 1989) assesses sleep quality and disturbances in the month before starting testing: in my experience with data collection online, participants have very different schedules and it is important to monitor that unusual lifestyles are not associated with sleep disturbances. The *Quality of Life Questionnaire* (QoL, Skevington et al., 2004) is a subjective measure of health and physical and social well-being, and it may impact cognition. *Positive and Negative Affect Schedule* (PANAS, Watson et al., 1988) is a self-rated mood scale, as mood can modulate training performance.

Working Memory Training

Intervention groups. Adaptive versions of the outcome DMTS and NB were used during the training sessions, with difficulty changing according to individual performance, through a staircase-like procedure (Buschkuehl et al., 2017; Truong et al., 2022), as follows. For the DMTS group, each session started with the easiest memory load condition (i.e., one colored target): if the participant reached 90% accuracy in a block, difficulty would be increased (i.e., by adding one target per block). Maximum difficulty was 6 targets, as visuo-spatial WM capacity in this task usually plateaus at around 3/4 elements (Vogel & Machizawa, 2004). Similarly, for the NB group, each session started with the easiest level ($n=1$) and increased in difficulty by one if 90% accuracy was reached in a block. Each training session comprised a total of 20 blocks of 20+n trials each (Jaeggi et al., 2010a, b; S. M. Jaeggi et al., 2008; Verhaeghen & Basak, 2005).

Control group. The control group served as an active control. During each training session, participants completed questionnaires without engaging in any formal cognitive training. This procedure matched the intervention groups in terms of online session scheduling and connection to the Prolific platform, ensuring equivalence in interaction across groups.

Statistical Analysis

All analyses were conducted with R (R Core Team, 2025). A full list of the packages used is provided in the **Supplementary Information**. This study's design and its analysis were not pre-registered. I report how I determined our sample size (see *Sample size and power calculations*), all data exclusions (if any), all manipulations, and all measures in the study, and the study follows JARS (Appelbaum et al., 2018). All data, analysis code, and research materials are available on the Open Science Framework at doi:https://doi.org/10.17605/OSF.IO/A3R6G. Only participants who completed the sessions within the prespecified times were included in the analysis.

Data cleaning and outlier detection. Individuals who did not complete all three outcome cognitive tasks at all time points (pre, post), with accuracy or reading speed above or below the mean by three standard deviations, or below chance in any of the tasks, were removed from further analysis. Missing data in questionnaires were imputed to their corresponding population mean (young or old), depending on the analysis (the whole population, for analysis comparing age groups, and the mean of each age group, for analysis within the group (Woods et al., 2023).

Demographic differences. T-tests were used to assess differences between groups.

Age-related effects at baseline. Linear regression models were used to predict the memory indexes (K_{st}, K_{up}, EMQ score, and reading comprehension %Corr) with age group, and years of formal education as fixed effects. In the case of reading comprehension, condition (fact-based, inferential) and reading speed were also included as fixed effects, to capture individual differences in processing efficiency that might influence reading comprehension. A Pearson correlation coefficient was computed to assess the linear relationship between working memory indices.

Training-related changes in performance. To test training-related effects on the outcome, I used linear mixed models to predict working memory performance (K_{st}, K_{up}) and reading comprehension performance (%Corr: 'fact-based', 'inferential'), including individuals as random factors. In all models, I used planned contrasts (DMTS versus CONTROL, NB versus CONTROL). I did not aim to directly compare the two training interventions (DMTS vs. NB), as my primary goal was to examine improvements within each training type and their potential transfer to reading comprehension. Assessing improvements in working memory performance within each intervention was necessary before evaluating transfer, as only demonstrated gains in the targeted cognitive domain could plausibly influence reading outcomes. Direct comparisons between interventions were not the focus, considering the distinct cognitive processes targeted by each approach. Including both younger and older adults was important to investigate age-related differences in baseline performance and transfer effects, but I did not explicitly compare training gains across age groups because the interventions were not designed or powered to test age × training interactions, and because the mechanisms underlying learning and transfer may differ between younger and older adults.

K_{st} was modelled by Session (pre, post), Training (CONTROL, DMTS, NBACK), and their interaction, and years of formal education. Because the n-Back task involves not only manipulation/updating but also storage of information, analyses on working memory updating (Kup) were also adjusted for baseline differences in storage capacity (Kst). This adjustment was made because updating performance depends on the amount of information that can be maintained in working memory, and controlling for Kst ensures that observed training effects on updating reflect genuine improvements in the updating component rather than differences in storage capacity.

Finally, the two reading conditions (fact-based, inferential) were modelled by adding to the model above both storage and updating working memory capacity at baseline ($K_{st,\_bas}$, K_{up_bas}) and reading speed, as these covariates capture individual differences in core cognitive resources and processing efficiency that are known to influence reading comprehension performance.

Relation between reading comprehension and working memory. All analyses were performed separately for each reading condition (fact-based, inferential). Individuals may rely on different mechanisms in completing the reading comprehension task, depending on their age or formal education, thus, performance was modeled separately for various age groups and adjusted for education, self-reported everyday memory, and reading speed to allow for the detection of differences in the dependencies. Linear regression models were computed to predict the performance at baseline for 'fact-based' and 'inferential' responses with working memory storage (K_{st}), updating (K_{up}) capacity, years of formal education, self-reported everyday memory, and reading speed as fixed effects. EMQ was included only in the reading comprehension models because it reflects self-perceived everyday memory rather than objective cognitive performance. Its inclusion allows me to account for potential influences of subjective memory experiences on reading comprehension, while in working memory models, which focus on objective measures, EMQ is not relevant.

Changes in reading performance were calculated as differences between post-training and pre-training performance and modelled with changes in working memory storage and updated capacity, years of education, and reading speed as fixed factors. I did not include EMQ in this model, as the focus here was on the relation between *training-related improvements* in working memory and reading comprehension. The EMQ reflects stable, self-reported everyday memory difficulties rather than session-to-session change and thus is not directly informative for testing whether gains in working memory predict corresponding gains in reading performance.

In all cases, model assumptions were evaluated through visual inspection of residual plots for normality and homoscedasticity. When necessary, follow-up analyses were conducted through planned contrasts, and Holm-adjusted for multiple comparisons. To ensure that including multiple covariates did not bias the results, multicollinearity was assessed using variance inflation factors (VIFs), which confirmed that all predictors were within acceptable limits (see Table S12).

Reporting. 95% confidence intervals (CI) are reported to aid data interpretation. Descriptive statistics is reported with standard error and confidence intervals and partial eta squared as effect sizes. Full model tables for each analysis, and tables containing Bayes factor, when available, are presented in the **Supplementary Information**. Bayes factors were reported in the main text to aid interpretation, with BF_{10} indicating the strength of evidence for the model including the factor (or interaction) of interest compared to a model excluding that factor.

Sample Size and Power Calculations

To estimate the required sample sizes, I modeled a multiple regression with a varying number of predictors (ranging from 2 to 7, covering the complexity of the models used here). Regarding effect size, I drew from meta-analyses in related domains as direct estimates for WM training and transfer to reading are scarce. A meta-analysis of 224 articles in memory research reported a median effect size of $\eta^2 = 0.18$ (Morris & Fritz, 2013). Additionally, a review in clinical psychology found a mean effect size for within-subject designs of $d = 0.75$, corresponding to $\eta^2 = 0.12$ for standardized mean changes (Rubio-Aparicio et al., 2018). While these sources are not specific to the present training-transfer context, they provide reasonable benchmarks for expected medium-to-large effects in cognitive intervention studies. Based on these findings, I conducted a power analysis for two possible effect sizes ($\eta^2_{medium} = 0.13$, $\eta^2_{large} = 0.26$), using a significance level of 0.05 and a power level of 0.80, implemented with the "pwr" package in R (Champely, 2020). For the most complex model, detecting a medium effect size would require a sample of 103 individuals, while a large effect size would require only 49 participants. When the focus of the study is on estimating the influence of fixed effects in a mixed model, and the random effects structure is relatively simple (e.g., single-level), power analyses based on multiple regression can provide reasonable approximations (Scherbaum & Ferreter, 2009).

Because the study was conducted fully online, I anticipated greater variability in performance compared to in-lab settings. For this reason, my recruitment strategy aimed toward a slightly larger sample. In practice, the final sample size was determined by the number of eligible participants that could be recruited within the time and budget constraints of the project. Data collection was stopped once these constraints were reached, rather than upon achieving the exact target suggested by the power analysis.

Results

Demographic characteristics by age are reported in Table 1. I found no baseline differences in education, familiarity with technology, or sleep quality between young and older adults (Table 2). I found a significant difference in overall quality of life, with young adults reporting worse quality of life than older adults (estimate = 1.37, CI[0.73, 2.02], t(173) = 4.23, p_{adj}<0.001, d_{Cohen}=0.64), a less positive attitude (estimate = 2.72, CI[0.41,5.03], t(173) = 2.33, p_{adj}=0.085, d_{Cohen}= 0.35) and a more negative attitude (estimate=−7.05, CI[−9.12,−4.99], t(173)= −6.763, p_{adj}<0.001, d_{Cohen}=−1.04).

Age-related Effects on Performance at Baseline

Working memory and everyday memory performance. Memory performance in young and older adults was quantified at baseline using objective (K_{st}, K_{up}) and subjective (EMQ score) measures. K_{st} is a measure of working memory storage capacity, while K_{up} is a proxy for working memory processing/updating capacity. Full model tables are reported in the Supplementary Information, Table S3. I found that age significantly predicted performance in all measures (K_{st}: b = 0.57, 95%CI [0.40;0.75], p < 0.001, η^2_p =0.20, BF_{10}=7.09*10^6; K_{up}: b=0.16, 95%CI [0.01;0.31], p=0.037, η^2_p= 0.03, BF_{10}=1.26; EMQ: b=21.28, 95%CI [12.83;29.73], p<0.001, η^2_p= 0.13, BF_{10}=9.36*10^3), with young individuals outperforming older adults in objective memory but older adults reporting better subjective performance that young adults. Education did not predict any of the memory measures. Correlational analysis showed that K_{st} and K_{up} are highly positively correlated both in young (r(81) = 0.34, p=0.002, 95%CI [0.14, 0.52]) and older adults (r(90)=0.28, p = 0.008, 95%CI [0.08, 0.46]). The subjective measure (EMQ) did not correlate with objective measures in young adults but did positively correlate with working memory updating capacity in older adults (K_{up}: r(90)=0.22, p = 0.035, 95%CI [0.02, 0.41]). While these sample sizes (≈80–90 per group) are sufficient for detecting medium size correlations, it should be noted that correlations generally stabilize slowly, and thus effect size estimates may still carry some uncertainty (Schönbrodt & Perugini, 2013).

Reading performance. Reading performance in young and older adults was quantified at baseline in terms of percentage of correct responses (%Corr) in facts-based and inferential questions. The full model tables are reported in the Supplementary Information, Table S4. I found a significant main effect of condition (b=−0.11, 95%CI [−0.13;−0.08], p < 0.001, η^2_p=0.28, BF_{10} = 1.17*10^{23}), with individuals being more accurate in facts-based questions than in inferential questions involving elaboration of the stored information. Reading speed also predicted the reading comprehension performance (b=0.14, 95%CI [0.07;0.20], p<0.001, η^2_p=0.04, BF_{10} = 312.63).

Training-related Changes in Performance

Descriptive statistics for each group are shown in Table 3**.** Full models are reported in the **Supplementary Information**: storage capacity in Table S5, updating capacity in Table S6, and reading comprehension in Table S7.

Changes in working memory storage capacity (K_{st}). I found a significant interaction of training per session, in both young and older adults (Young: b = 0.25, 95%CI [0.05;0.72], p = 0.025, η^2_p=0.06, B_{10} = 0.24; Old: b=0.21, 95%CI [0.03;0.40], p=0.022, η^2_p=0.07, BF_{10}=0.23), with K_{st} significantly improving with respect to control only in

Table 2 Comparison of demographic characteristics between age groups. Characteristics that are significantly different between age groups are marked in bold.

	OLD - YOUNG					
	ESTIMATE	95% CI		t(173)	p_{adj}	cohen's d
		LL	UL			
gender (f/m)*	--	--	--	--	0.287	--
education (yrs)	−0.78	−1.67	0.11	−1.74	0.253	−0.26
Tech	−0.14	−1.14	0.85	−0.29	1.000	−0.04
QoL	1.37	0.73	2.02	4.23	**<0.001**	0.64
Positive affect	2.72	0.41	5.03	2.33	0.085	0.35
Negative affect	−7.05	−9.12	−4.99	−6.76	**<0.001**	−1.04
PSQI	−0.23	−1.01	0.56	−0.57	1.000	−0.09
EMQ	−21.07	−29.61	−12.53	−4.88	**<0.001**	−0.75

* Chi-square test X(n = 175, df = 1) = 1.13

Abbreviations: Tech= Familiarity with Technology, QoL= Quality of Life, PSQI= Pittsburgh Sleep Quality Index, EMQ= Everyday Memory Questionnaire, Positive affect = PANAS Positive Affect score, Negative affect= PANAS Negative Affect score

Table 3 Descriptive statistics in each training group. ΔKst = post-pre change in working memory storage capacity; ΔKup = post-pre change in working memory updating capacity; Δ%Corr = post-pre change in percentage of correct responses in reading comprehension

YOUNG	CONTROL				DMTS				NB			
	M	SE	95% CI		M	SE	95% CI		M	SE	95% CI	
			LL	UL			LL	UL			LL	UL
age	24.20	0.51	23.17	25.23	24.40	0.68	23.00	25.80	23.90	0.71	22.45	25.35
education (yrs)	16.70	0.36	15.96	17.44	16.30	0.45	15.38	17.22	15.10	0.54	13.99	16.21
ΔKst	−0.04	0.12	−0.28	0.20	0.35	0.12	0.10	0.60	0.11	0.13	−0.16	0.37
ΔKpr	0.09	0.08	−0.07	0.25	0.11	0.07	−0.04	0.26	0.51	0.13	0.24	0.79
Δ%Corr	0.00	0.01	−0.03	0.03	−0.04	0.03	−0.09	0.02	0.01	0.02	−0.03	0.05
OLD												
	CONTROL				DMTS				NB			
	M	SE	95% CI		M	SE	95% CI		M	SE	95% CI	
			LL	UL			LL	UL			LL	UL
age	67.80	0.51	66.75	68.85	67.50	0.52	66.43	68.57	68.60	0.52	67.55	69.65
education (yrs)	15.60	0.65	14.27	16.93	15.00	0.67	13.63	16.37	15.30	0.60	14.08	16.52
ΔKst	0.17	0.04	0.09	0.26	0.39	0.09	0.20	0.58	0.19	0.05	0.08	0.30
ΔKpr	0.12	0.04	0.04	0.20	0.08	0.04	−0.01	0.16	0.28	0.06	0.17	0.40
Δ%Corr	0.02	0.02	−0.01	0.05	0.00	0.01	−0.02	0.03	0.01	0.01	−0.01	0.03

the groups receiving DMTS training. Individual performance is shown in Fig. 3, panel A.

Changes in working memory updating capacity (K_{up}). After adjusting for baseline differences in working memory storage capacity (K_{st}), I found a significant interaction of training per session, in both young and older adults (Young: b=0.42, 95%CI [0.16;0.69], p=0.002, η_p^2=0.13, BF_{10}=0.37; Old: b=0.16, 95%CI [0.03;0.29], p=0.016, η_p^2=0.11, BF_{10}=0.36), with K_{up} significantly improving with respect to control only in the groups receiving n-back training. Individual performance is shown in Fig. 3, panel B.

Changes in reading comprehension performance. I found no significant effects of training working memory on reading comprehension performance, either in the fact-based or inferential condition. Individual performance is shown in Fig. 4. Thus, training produced statistically significant improvements in the working memory component targeted by each training task, as indicated by the frequentist analyses (K_{st} with DMTS training and K_{up} with NB training) in young and older adults, which is in line with previous studies (Waris et al., 2015), but no changes in reading comprehension were found at the group level. However, the associated Bayes factors provided only weak evidence for these effects, suggesting that the magnitude and robustness of the training-related gains should be interpreted cautiously.

Relation between Reading Comprehension and Working Memory Capacity

I performed a planned analysis to understand whether individuals who improved after training in either working memory storage (K_{st}) or process (K_{up}) capacity also improved in reading comprehension, regardless of the type of training received.

At baseline, before any training took place, I found no effect of working memory storage capacity K_{st} on reading performance, regardless of the question type (fact-based, inferential) or the age group considered. Working memory updating capacity K_{up} significantly predicted performance for fact-based questions in older adults only (b=−0.04, 95%CI [−0.08;−0.01], p=0.017, η_p^2=0.04). However, Bayes factors provided only anecdotal evidence in either direction (inclusion or exclusion of K_{st} and K_{up} from the model, all BFs<=3). Table S8 and Table S9 in **Supplementary Information** report the full models and BF01 and BF10 for all models and conditions, respectively.

When modeling training-related changes in reading performance (fact-based or inferential) as a function of changes in working memory storage (K_{st}) or updating (K_{up}) capacity, while adjusting for education and reading speed, I found that changes in K_{st} significantly predicted same-direction

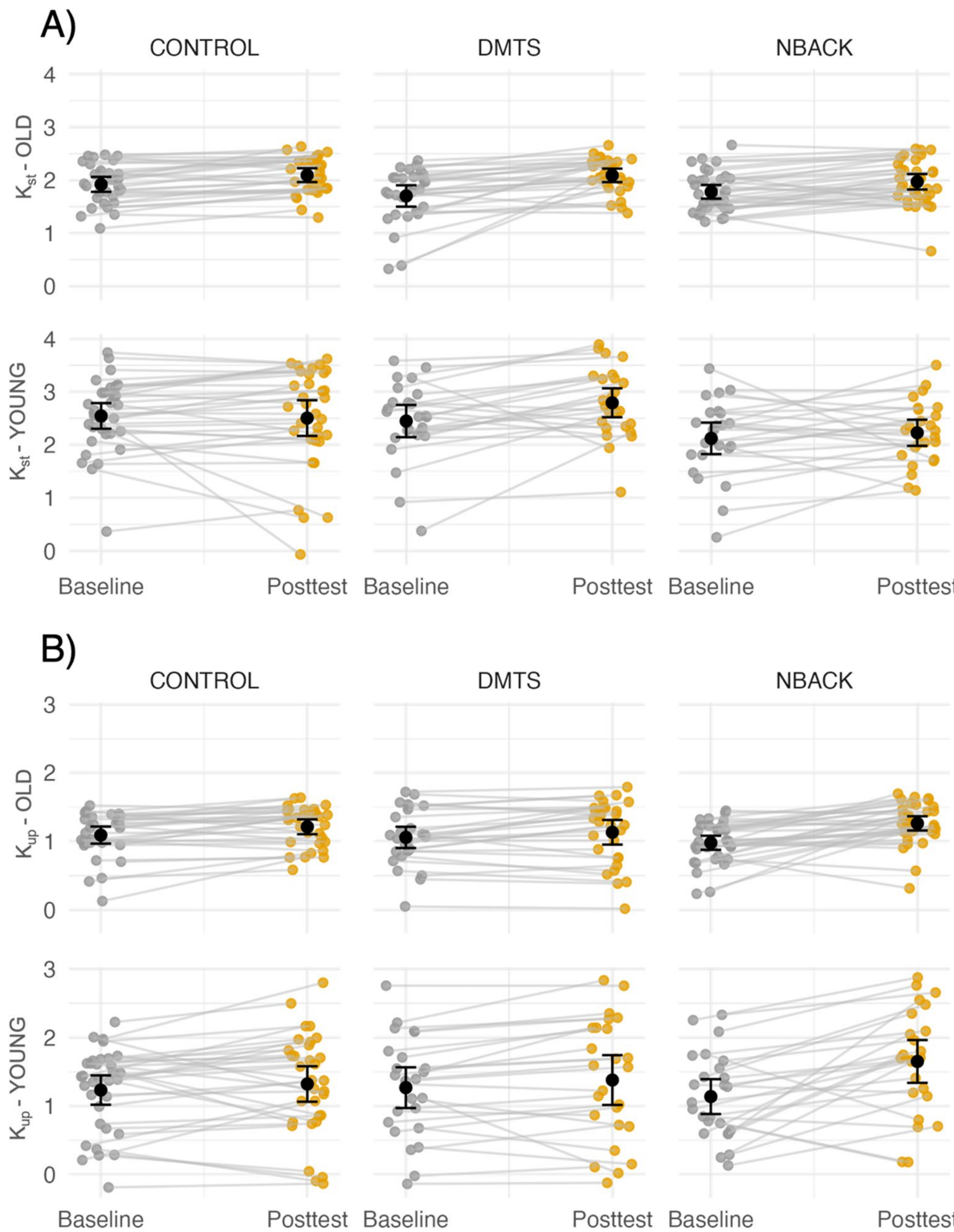

Fig. 3 Working memory storage (*Kst*, panel A) and updating (*Kup*, panel B) capacity in young and older adults, before and after training (CONTROL, DMTS, NBACK). Values represent raw means

changes in inferential performance in young adults only ($b=0.05$, 95%CI [0.01;0.09], $p=0.013$, $\eta_p^2=0.06$), with the associated Bayes factor providing moderate evidence for the relationship ($BF_{10}=4.30$). Details of the models, and corresponding Bayes factors are reported in the **Supplementary Information**, Table S10, Table S11.

Discussion

The primary objective of this study was to investigate the relationship between memory (storage capacity - K_{st}, updating capacity - K_{up}, and Everyday Memory Questionnaire - EMQ) and reading comprehension performance (as a proxy for everyday activities), and their sensitivity to age differences. Additionally, I examined whether training-related changes in working memory transfer to enhancements in a reading comprehension task.

In this study, I chose to investigate these questions using a brief working memory training program focused on visuo-spatial memory. While this choice may seem counterintuitive, it was informed by previous findings from our lab, where visuo-spatial adaptive tasks have consistently led to improvements after just a few sessions (Assecondi et al., 2022; Assecondi, Villa-Sánchez et al., 2022; Tagliabue et al., 2024). A key question remains whether such improvements

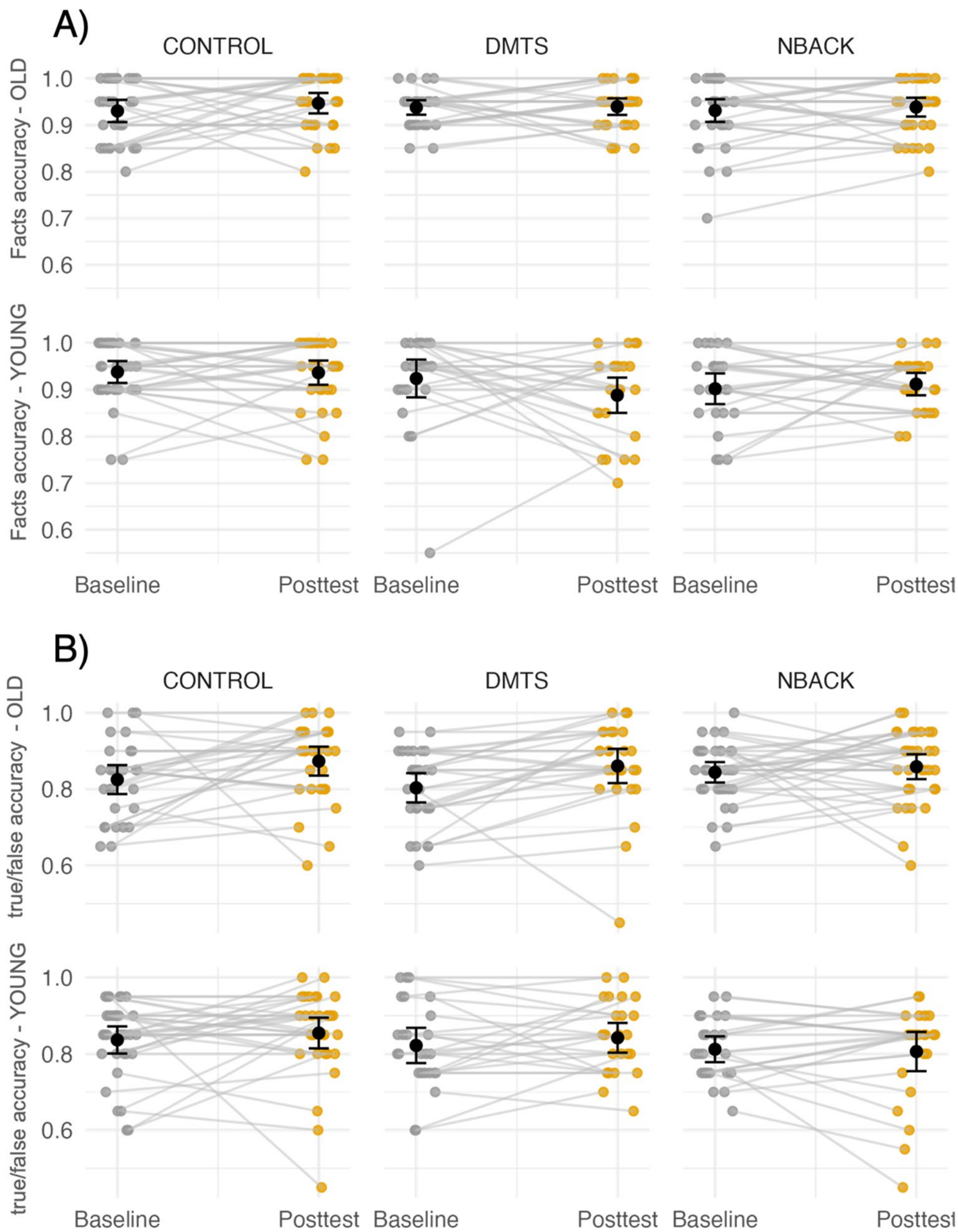

Fig. 4 Reading comprehension performance, expressed as the percentage of correct responses, for "fact-based" (panel A) and "inferential" (panel B) questions in young and older adults, before and after training. Values represent raw means

extend beyond trained tasks and transfer to a more ecologically relevant ability, such as reading comprehension. This consideration aligns with the broader debate on domain-general versus domain-specific mechanisms in cognitive training (Peng et al., 2018).

The decision to focus on visuo-spatial working memory rather than verbal working memory was further motivated by several advantages of visuo-spatial training. First, visuo-spatial tasks reduce linguistic dependency, making them more resilient to biases associated with cultural differences in language, idioms, and educational backgrounds, thus enhancing their generalizability. Second, visuo-spatial tasks are inherently more engaging and dynamic, which can improve participants' motivation and compliance. This is particularly relevant for home-based online experiments, where participant attrition is a concern.

Age Differences in Memory Performance

In this study, I assessed working memory capacity using two different indices: K_{st}, derived from a delayed match-to-sample task, which primarily reflects storage capacity, and K_{up}, obtained from an n-back task, which captures both storage and updating processes. Additionally, I included a self-report measure (EMQ) to evaluate perceived memory performance at baseline.

Consistent with my first hypothesis, young adults significantly outperformed older adults in objective memory tasks, confirming well-established age-related declines in working memory storage and updating capacity (Bopp & Verhaeghen, 2020). However, this hypothesis was not fully supported for subjective memory performance, as older adults reported fewer subjective memory failures than younger adults. This discrepancy between objective and self-perceived memory performance may reflect age-related differences in metacognitive awareness, memory expectations, or compensatory strategies rather than actual functional abilities.

Interestingly, I found no significant age-related differences in reading comprehension performance, aligning with previous studies (Cansino et al., 2013; Noort et al., 2006): while working memory declined with age, reading comprehension remained relatively preserved. However, question type played a key role, with participants performing better on factual questions than on inferential questions requiring elaborative processing. This suggests that more complex comprehension tasks engage additional cognitive demands, consistent with prior research on reading comprehension difficulty (Schroeder, 2014).

One limitation of this study is the use of a naturalistic reading comprehension task. While such tasks have ecological validity by better reflecting real-world reading experiences, they may lack the structured format of standardized assessments like the Nelson-Denny reading test (Brown et al., 1993), potentially reducing sensitivity to subtle differences in comprehension performance. This trade-off between ecological validity and experimental control should be considered when interpreting the findings.

Greater variability was observed among younger adults, including some individuals who performed worse at post-test (Figs. 3 and 4). This may reflect the online testing modality, as younger participants are more likely to multitask or be distracted during web-based tasks (Kostić & Ranđelović, 2022). Older adults, by contrast, may approach the tasks with greater motivation and commitment, resulting in more stable performance. Recent evidence also suggests that in online cognitive research, *who* is tested is often more consequential than *how* the testing is conducted (Uittenhove et al., 2023), which could further explain the observed differences in variability across age groups. Such factors should be considered when interpreting individual differences in training outcomes and transfer effects.

To summarise, while age-related declines were evident in both storage and updating capacities, reading comprehension performance remained relatively stable across groups. This suggests that reading comprehension may draw not only on core working memory resources but also on compensatory mechanisms such as accumulated linguistic knowledge, contextual inference, or strategy use, especially in older adults. Moreover, the stronger cognitive demands of inferential questions highlight that the extent to which working memory supports comprehension depends on the type of task involved. Finally, the discrepancy between objective and subjective memory assessments underscores the importance of using objective measures when investigating cognitive predictors of comprehension, as self-reported memory complaints may be shaped more by metacognitive factors than by actual capacity. Together, these results indicate that working memory plays a role in reading comprehension, but its contribution is modulated by age, task demands, and compensatory processes.

Overall Training Effects

The present study supports the notion that training effects in working memory are process-specific. Consistent with my hypotheses, training led to improvements in the targeted processes: working memory storage capacity (K_{st}) increased following DMTS training and working memory updating capacity (K_{up}) improved following n-back training, in both age groups, relative to the control condition. Although the associated Bayes factors indicated only weak evidence, this may reflect the relatively brief training duration rather than the absence of genuine effects. Overall, these results suggest that even short visuo-spatial training interventions can produce selective and meaningful gains in the intended working memory components.

In line with previous findings (e.g., Matysiak et al., 2019) the efficacy of n-back training appears to be moderated by individuals' initial capacity, suggesting that baseline ability may influence the extent to which training-related benefits can be realized. This aligns with the view that training outcomes in working memory are constrained by individual differences, which has implications for tailoring interventions.

However, despite these training-related gains in working memory capacity, the hypothesis that such improvements would transfer to reading comprehension was not supported at the group level. This finding point to a limited generalizability of working memory training effects and suggests that transfer to complex tasks like reading comprehension may require broader or more integrative training approaches.

A limitation of the current study is the nature of the control group. Although control participants completed questionnaires to match the training schedule of the intervention groups, they did not engage in an alternative cognitive task. Consequently, placebo effects, expectancy, or engagement-related influences cannot be fully ruled out, which should be considered when interpreting the observed training-related improvements.

Relation between Working Memory Capacity and Reading Comprehension Performance

Age-related differences emerged when considering the relationship between reading comprehension performance and working memory capacity. At baseline, our results indicated that working memory storage capacity (K_{st}) did not predict reading comprehension performance, irrespective of question type (fact-based or inferential) or age group, while working memory updating capacity (K_{up}) significantly predicted performance for fact-based questions in older adults only. Thus, our hypothesis that individual differences in reading comprehension would be explained by working memory capacity was not broadly supported, with only limited and age-specific evidence for an association between K_{up} and factual comprehension in older adults. It should also be noted that the observed effect of updating (K_{up}) on reading comprehension emerged only when self-reported memory complaints (EMQ) were included as a covariate, and the associated Bayes factor provided only anecdotal evidence, warranting cautious interpretation.

These findings suggest that, overall, there is no robust baseline relationship between objective working memory measures and reading comprehension performance in a naturalistic task. In older adults, however, the, albeit weak, association between K_{up} and fact-based reading performance may reflect the multifaceted nature of the n-back task, which, although designed to assess both storage and updating, appears to tap more into attentional, verbal memory, and updating processes in this age group (Gajewski et al., 2018). Given that reading comprehension involves several cognitive steps, from encoding new information to retrieving and integrating previously stored knowledge (Gordon et al., 2016), it is plausible that in older adults, K_{up} is a better predictor of performance when these processes are engaged.

Furthermore, these results resonate with previous meta-analytic findings (McVay & Kane, 2012) suggesting that working memory capacity alone, especially as measured by simple span tasks, is less predictive of reading comprehension than more complex tasks that incorporate strategy use (see also Carretti et al., 2009; Engle et al., 1999). Thus, the link between working memory and reading comprehension cannot be reduced to raw storage or updating measures, but instead depends on the interplay between task demands and individual strategy use. It is possible that the variability in cognitive strategies, particularly in older adults, who may rely more on accumulated knowledge to offset declines in raw memory capacity (Hannon & Daneman, 2009), obscures a straightforward relationship between capacity measures (K_{st} or K_{up}) and reading outcomes (Bailey et al., 2008; Dunlosky & Kane, 2007; Friedman & Miyake, 2004).

Lastly, the lack of a significant relationship between the self-reported Everyday Memory Questionnaire (EMQ) scores and reading performance in both age groups may indicate that subjective assessments of memory are prone to biases or do not accurately capture the specific cognitive processes involved in reading comprehension (Sunderland et al., 1983; Zelinski et al., 1980). Given that the EMQ captures subjective memory complaints rather than direct cognitive performance, its lack of association with reading comprehension is not unexpected. Although the EMQ has been validated in healthy adults, differences in self-awareness or metacognitive monitoring across age groups may still influence the relationship between subjective memory complaints and objective working memory performance. This should be considered when interpreting the observed correlations in younger versus older participants.

Improving Reading Comprehension Performance by Training Working Memory

When examining training-related changes, I found that improvements in working memory storage capacity (K_{st}) significantly predicted enhancements in reading comprehension, specifically for inferential questions, in young adults only, with moderate evidence supporting this relationship ($BF_{10}=4.30$). Thus, our hypothesis that training-related gains in working memory would transfer to reading comprehension was only partially supported: transfer effects emerged descriptively in younger adults but were absent in older adults. Although no formal age-group comparisons were conducted, the pattern of results suggests that the extent to which working memory improvements transfer to reading comprehension may differ descriptively across age groups, with preliminary evidence of a modest relationship in younger but not older adults. Importantly, however, the Bayesian evidence was only moderate, and so this finding should be interpreted with caution rather than as a definitive effect.

This finding tentatively suggests that in young adults, gains in working memory storage capacity may facilitate more effective processing in tasks that require elaborative integration of information. In contrast, no such relationship was observed in older adults, implying that for this group, increases in working memory capacity alone may not be sufficient to boost reading comprehension. Older adults might instead rely on compensatory strategies or prior knowledge to support performance on complex comprehension tasks. Taken together, these results indicate that our transfer hypothesis was not broadly supported, but they point to potential age-related differences in the mechanisms underlying transfer.

It should be noted that although performance on the fact-based questions was generally high, suggesting a potential ceiling effect, there was still some variability across participants. This limitation may have constrained the sensitivity to detect transfer effects of working memory training on fact-based comprehension, potentially contributing to the lack of significant improvements observed for these items, despite clear gains in working memory capacity.

Furthermore, in our data, baseline working memory and reading comprehension were not strongly or reliably correlated. At first glance, this might raise the question of what a "successful" transfer effect would imply under such circumstances. However, the absence of a robust cross-sectional association at baseline does not preclude the possibility of training-related changes leading to transfer. I suggest that transfer reflects a dynamic process rather than a static correlation: improvements in working memory may change how readers engage with texts, even if baseline associations are weak. For example, enhanced working memory may support better allocation of attentional resources, more efficient integration of information across sentences, and greater flexibility in updating situational models during reading (Martin et al., 2020; Daneman & Merikle, 1996). These are mechanisms that would not necessarily be reflected in a simple baseline correlation but may emerge when cognitive capacity is experimentally enhanced. Thus, evidence of transfer should be interpreted as reflecting changes in functional engagement of cognitive resources rather than a mere amplification of baseline associations.

One potential explanation for the limited transfer is that our training focused on visuo-spatial working memory, which may not engage the verbal-processing mechanisms most critical for reading performance. Previous studies have shown that while visual working memory is associated with reading comprehension (e.g., Pham & Hasson, 2014; Walczyk & Taylor, 1996), in a complex way. For example, while general working memory correlates with sentence comprehension (Moser et al., 2007), visual working memory specifically predicts reading comprehension but not necessarily reading fluency (Pham & Hasson, 2014) and its predictive power may differ from that of verbal complex span tasks (Daneman & Merikle, 1996; Kane et al., 2004). Thus, our findings suggest that to achieve stronger and more consistent transfer effects on reading abilities, future interventions may need to incorporate training paradigms that engage verbal as well as visuo-spatial working memory processes.

To conclude, the present study shows that working memory training produced process-specific improvements, but transfer to reading comprehension was modest. Gains in storage capacity predicted enhanced comprehension in young adults under certain conditions, whereas no such effects were observed in older adults. For young adults, gains in storage capacity supported performance on more demanding comprehension tasks, whereas for older adults, such improvements did not translate into reading benefits, pointing to the need for additional or alternative approaches. These results highlight the nuanced relationship between working memory and reading comprehension and suggest that future interventions may benefit from multidomain training programs that combine different cognitive processes and account for age-specific strategies to promote more meaningful transfer to real-world skills.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s41465-025-00342-4.

Acknowledgements We thank research assistants Cecilia Barlabà, Eleonora Malloggi, Nicoletta Romano, and Greta Varesio for their assistance in data collection. We are also grateful to Elena Scagliotti and Fausta Onofrio for their help in gathering stimuli for the reading comprehension task. Additionally, we thank Chiara Tagliabue for insightful discussions and Veronica Mazza for both valuable discussions and constructive feedback on an earlier version of the manuscript.

Author Contributions SA served as lead for data curation, formal analysis, investigation, methodology, project administration, software, supervision, validation, visualization, and writing.

Funding Open access funding provided by Università degli Studi di Trento within the CRUI-CARE Agreement. SA is supported by Fondazione Cassa di Risparmio di Trento e Rovereto (CARITRO) [grant number 000040103444CARITRO].

Data Availability Data and study materials are available on the Open Science Framework (doi:10.17605/OSF.IO/A3R6G). The study was not preregistered.

Declarations

Ethics Approval The study was approved by the University of Trento Research Ethics Committee (Protocol No. 2021-041).

Informed Consent All participants provided informed consent before taking part in the study, in accordance with ethical guidelines for research involving human subjects.

Competing Interests SA is named inventor on a patent application (publication number WO/2022/106850) jointly submitted by the University of Birmingham and Dalhousie University, titled "Improving cognitive function," (international application No. PCT/G82021/053019).

References

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

Assecondi, S., Hu, R., Eskes, G., Pan, X., Zhou, J., & Shapiro, K. (2021). Impact of tDCS on working memory training is enhanced by strategy instructions in individuals with low working memory capacity. *Scientific Reports, 11*(1), Article 1. https://doi.org/10.1038/s41598-021-84298-3

Assecondi, S., Hu, R., Kroeker, J., Eskes, G., & Shapiro, K. (2022). Older adults with lower working memory capacity benefit from transcranial direct current stimulation when combined with working memory training: A preliminary study. *Frontiers in Aging Neuroscience*. https://doi.org/10.3389/fnagi.2022.1009262

Assecondi, S., Villa-Sánchez, B., & Shapiro, K. (2022). Event-related potentials as markers of efficacy for combined working memory training and transcranial direct current stimulation regimens: A proof-of-concept study. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/fnsys.2022.837979

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition*, *36*(8), 1383–1390. https://doi.org/10.3758/mc.36.8.1383

Bissig, D., Kaye, J., & Erten-Lyons, D. (2020). Validation of SATURN, a free, electronic, self-administered cognitive screening test. *Alzheimer's & Dementia : Translational Research & Clinical Interventions, 6*(1), Article e12116. https://doi.org/10.1002/trc2.12116

Bopp, K. L., & Verhaeghen, P. (2020). Aging and n-Back performance: A meta-analysis. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 75*(2), 229–240. https://doi.org/10.1093/geronb/gby024

Borella, E., Carbone, E., Pastore, M., Beni, R. D., & Carretti, B. (2017). Working memory training for healthy older adults: The role of individual characteristics in explaining short- and long-term gains. *Frontiers in Human Neuroscience, 11*, 99. https://doi.org/10.3389/fnhum.2017.00099

Brown, J. A., Fishco, V. V., & Hanna, G. S. (1993). *Nelson–Denny Reading Test: Manual for Scoring and Interpretation* (Forms G&H. Rolling Meadows,). IL: Riverside Publishing.

Buschkuehl, M., Jaeggi, S. M., Mueller, S. T., Shah, P., & Jonides, J. (2017). Training change detection leads to substantial task-specific improvement. *Journal of Cognitive Enhancement, 1*(4), 419–433. https://doi.org/10.1007/s41465-017-0055-y

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Cansino, S., Hernández-Ramos, E., Estrada-Manilla, C., Torres-Trejo, F., Martínez-Galindo, J. G., Ayala-Hernández, M., Gómez-Fernández, T., Osorio, D., Cedillo-Tinoco, M., Garcés-Flores, L., Beltrán-Palacios, K., García-Lázaro, H. G., García-Gutiérrez, F., Cadena-Arenas, Y., Fernández-Apan, L., Bärtschi, A., & Rodríguez-Ortiz, M. D. (2013). The decline of verbal and visuospatial working memory across the adult life span. *Age (Dordrecht, Netherlands), 35*(6), 2283–2302. https://doi.org/10.1007/s11357-013-9531-1

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin & Review*, *20*(2), 243–268. https://doi.org/10.3758/s13423-012-0369-9

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, *19*(2), 246–251. https://doi.org/10.1016/j.lindif.2008.10.002

Carretti, B., Caldarola, N., Tencati, C., & Cornoldi, C. (2014). Improving reading comprehension in reading and listening settings: The effect of two training programmes focusing on metacognition and working memory. *British Journal of Educational Psychology, 84*, 194–210. https://doi.org/10.1111/bjep.12022

Champely, S. (2020). *pwr: Basic Functions for Power Analysis* (Version R package version 1.3-0) [Computer software]. https://CRAN.R-project.org/package=pwr

Cohen, R. A., Marsiske, M. M., & Smith, G. E. (2019). Chapter 10—Neuropsychology of aging. In S. T. Dekosky & S. Asthana (Eds.), *Handbook of Clinical Neurology* (Vol. 167, pp. 149–180). Elsevier. https://doi.org/10.1016/B978-0-12-804766-8.00010-8

R Core Team (2025). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/s0140525x01003922

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433. https://doi.org/10.3758/BF03214546

De Beni, R., Palladino, P., Borella, E., & Lo Presti, S. (2003). Reading comprehension and aging: Does an age-related difference necessarily mean impairment? *Aging Clinical and Experimental Research*, *15*(1), 67–76. https://doi.org/10.1007/BF03324482

De Beni, R., Borella, E., & Carretti, B. (2007). Reading comprehension in aging: The role of working memory and metacomprehension. *Aging Neuropsychology and Cognition*, *14*(2), 189–212. https://doi.org/10.1080/13825580500229213

DeCaro, R., Peelle, J. E., Grossman, M., & Wingfield, A. (2016). The Two Sides of Sensory–Cognitive Interactions: Effects of Age, Hearing Acuity, and Working Memory Span on Sentence Comprehension. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyg.2016.00236

Deveau, J., Jaeggi, S. M., Zordan, V., Phung, C., & Seitz, A. R. (2015). How to build better memory training games. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00243

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227–1245. https://doi.org/10.1080/17470210600926075

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W., & Kane, M. J. (2003). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. *Psychology of Learning and Motivation* (Vol. 44, pp. 145–199). Elsevier. https://doi.org/10.1016/S0079-7421(03)44005-X

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. R. A. (1999). *Working Memory, Short-Term Memory, and General*

Fluid Intelligence: A Latent-Variable Approach. 23. https://doi.org/10.1037/0096-3445/99/S3.00.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Fan, B. J. Y., & Wong, R. Y. M. (2019). Effect of cognitive training on daily function in older people without major neurocognitive disorder: A systematic review. *Geriatrics*, *4*(3), 44. https://doi.org/10.3390/geriatrics4030044

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language*, *51*(1), 136–158. https://doi.org/10.1016/j.jml.2004.03.008

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the n-Back task measure as we get older? Relations between Working-Memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology*, *9*, 2208. https://doi.org/10.3389/fpsyg.2018.02208

Gates, N. J., Vernooij, R. W., Nisio, M. D., Karim, S., March, E., Martínez, G., & Rutjes, A. W. (2019). Computerised cognitive training for preventing dementia in people with mild cognitive impairment. *Cochrane Database of Systematic Reviews*, *3*(3), CD012279. https://doi.org/10.1002/14651858.cd012279.pub2

Gobet, F., & Sala, G. (2023). Cognitive training: A field in search of a phenomenon. *Perspectives on Psychological Science*, *18*(1), 125–141. https://doi.org/10.1177/17456916221091830

Goghari, V. M., & Lawlor-Savage, L. (2018). Self-perceived benefits of cognitive training in healthy older adults. *Frontiers in Aging Neuroscience*. https://doi.org/10.3389/fnagi.2018.00112

Gordon, P. C., Lowder, M. W., & Hoedemaker, R. S. (2016). Reading in normally aging adults. In H. H. Wright (Ed.), *Cognition, Language and Aging* (pp. 165–191). John Benjamins Publishing Company. https://doi.org/10.1075/z.200.07gor

Hannon, B., & Daneman, M. (2009). Age-related changes in reading comprehension: An individual-differences perspective. *Experimental Aging Research, 35*(4), 432–456. https://doi.org/10.1080/03610730903175808

Hou, J., Jiang, T., Fu, J., Su, B., Wu, H., Sun, R., & Zhang, T. (2020). The long-term efficacy of working memory training in healthy older adults: A systematic review and meta-analysis of 22 randomized controlled trials. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 75*(8), e174–e188. https://doi.org/10.1093/geronb/gbaa077

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the *N* -back task as a working memory measure. *Memory, 18*(4), 394–412. https://doi.org/10.1080/09658211003702171

Jaeggi, S., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W. (2010). The relationship between n-back performance and matrix reasoning—Implications for training and transfer. *Intelligence, 38*, 625–635. https://doi.org/10.1016/j.intell.2010.09.001

Jayalath, J., & Esichaikul, V. (2022). Gamification to enhance motivation and engagement in blended eLearning for technical and vocational education and training. *Technology, Knowledge and Learning, 27*(1), 91–118. https://doi.org/10.1007/s10758-020-09466-2

Just, M. A., & Carpenter, P. A. (1992). *A Capacity Theory of Comprehension: Individual Differences in Working Memory*. 28.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General, 133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Katz, B., Shah, P., & Meyer, D. E. (2018). How to play 20 questions with nature and lose: Reflections on 100 years of brain-training research. *Proceedings of the National Academy of Sciences of the United States of America, 115*(40), 9897–9904. https://doi.org/10.1073/pnas.1617102114

Kim, D., Mishra, S., Wang, Z., & Singh, S. N. (2016). Insidious effects of syntactic complexity: Are ads targeting older adults too complex to remember? *Journal of Advertising, 45*(4), 509–518. https://doi.org/10.1080/00913367.2016.1262301

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Koivisto, J., & Malik, A. (2020). Gamification for older adults: A systematic literature review. *The Gerontologist*. https://doi.org/10.1093/geront/gnaa047

Kostić, J. O., & Ranđelović, K. R. (2022). *Digital distractions: learning in multitasking environment*. 301–304. https://doi.org/10.36315/2022inpact070

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a virtual chinrest. *Scientific Reports, 10*(1), 904. https://doi.org/10.1038/s41598-019-57204-1

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Martin, J. D., Shipstead, Z., Harrison, T. L., Redick, T. S., Bunting, M., & Engle, R. W. (2020). The role of maintenance and disengagement in predicting reading comprehension and vocabulary learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 46*(1), 140–154. https://doi.org/10.1037/xlm0000705

Martín-Aragoneses, M. T., Mejuto, G., Del Río, D., Fernandes, S. M., Rodrigues, P. F. S., & López-Higes, R. (2023). Task demands and sentence reading comprehension among healthy older adults: The complementary roles of cognitive reserve and working memory. *Brain Sciences, 13*(3), Article 428. https://doi.org/10.3390/brainsci13030428

Matysiak, O., Kroemeke, A., & Brzezicka, A. (2019). Working memory capacity as a predictor of cognitive training efficacy in the elderly population. *Frontiers in Aging Neuroscience*, *11*, 126. https://doi.org/10.3389/fnagi.2019.00126

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology. General, 141*(2), 302–320. https://doi.org/10.1037/a0025250

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291. https://doi.org/10.1037/a0028228

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of far transfer. *Perspectives on Psychological Science, 11*(4), 512–534. https://doi.org/10.1177/1745691616635612

Miller, K., & Price, C. & Clinical … O.-M. of. (2009). *Is the N-Back Task a Valid Neuropsychological Measure for Assessing Working Memory?*.

Morris, P. E., & Fritz, C. O. (2013). Effect sizes in memory research. *Memory, 21*(7), 832–842. https://doi.org/10.1080/09658211.2013.763984

Moser, D. C., Fridriksson, J., & Healy, E. W. (2007). Sentence comprehension and general working memory. *Clinical Linguistics & Phonetics*, *21*(2), 147–156. https://doi.org/10.1080/02699200600782526

Naveh-Benjamin, M., & Cowan, N. (2023). The roles of attention, executive function and knowledge in cognitive ageing of working memory. *Nature Reviews Psychology, 2*(3), Article 3. https://doi.org/10.1038/s44159-023-00149-0

Ng, N. F., Schafer, R. J., Simone, C. M., & Osman, A. M. (2020). Perceptions of brain training: Public expectations of cognitive benefits from popular activities. *Frontiers in Human Neuroscience*, *14*, 15. https://doi.org/10.3389/fnhum.2020.00015

Norman, S., Kemper, S., & Kynette, D. (1992). Adults' reading comprehension: Effects of syntactic complexity and working memory. *Journal of Gerontology*, *47*(4), P258–P265. https://doi.org/10.1093/geronj/47.4.P258

Osborne-Crowley, K. (2020). Social cognition in the real world: Reconnecting the study of social cognition with social reality. *Review of General Psychology*, *24*(2), 144–158. https://doi.org/10.1177/1089268020906483

Ossher, L., Flegal, K. E., & Lustig, C. (2013). Everyday memory errors in older adults. *Aging, Neuropsychology, and Cognition, 20*(2), 220–242. https://doi.org/10.1080/13825585.2012.690365

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

Payne, B. R., & Stine-Morrow, E. A. L. (2017). The effects of home-Based cognitive training on verbal working memory and language comprehension in older adulthood. *Frontiers in Aging Neuroscience, 9*, 256. https://doi.org/10.3389/fnagi.2017.00256

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, *144*(1), 48–76. https://doi.org/10.1037/bul0000124

Pham, A. V., & Hasson, R. M. (2014). Verbal and visuospatial working memory as predictors of children's reading ability. *Archives of Clinical Neuropsychology, 29*(5), 467–477. https://doi.org/10.1093/arclin/acu024

Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging*, *16*(1), 145–160. https://doi.org/10.1037/0882-7974.16.1.145

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., Marsiske, M., Morris, J. N., Tennstedt, S. L., Unverzagt, F. 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*(1), 16–24. https://doi.org/10.1111/jgs.12607

Redlinger, E., Glas, B., & Rong, Y. (2022). Impact of visual game-like features on cognitive performance in a virtual reality working memory task: Within-subjects experiment. *JMIR Serious Games, 10*(2), e35295. https://doi.org/10.2196/35295

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*(4), 813. https://doi.org/10.1037/a0023631

Rodrigues, P. F. S., Bártolo, A., Ribeiro, B., López-Higes, R., Rubio-Valdehita, S., Caetano, A. P., & Fernandes, S. M. (2025). Everyday memory questionnaire—Revised (EMQ-R): Psychometric validation of the European Portuguese version in non-clinical sample. *Behavioral Sciences, 15*(3), 280. https://doi.org/10.3390/bs15030280

Royle, J., & Lincoln, N. B. (2009). The everyday memory questionnaire – revised: Development of a 13-item scale. *Disability and Rehabilitation, 30*(2), 114–121. https://doi.org/10.1080/09638280701223876

Rubio-Aparicio, M., Marín-Martínez, F., Sánchez-Meca, J., & López-López, J. A. (2018). A methodological review of meta-analyses of the effectiveness of clinical psychology treatments. *Behavior Research Methods*, *50*(5), 2057–2073. https://doi.org/10.3758/s13428-017-0973-8

Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A Second-Order Meta-Analysis. *Collabra: Psychology, 5*(1), 18. https://doi.org/10.1525/collabra.203

Sander, M. C., Lindenberger, U., & Werkle-Bergner, M. (2012). Lifespan age differences in working memory: A two-component framework. *Neuroscience and Biobehavioral Reviews, 36*(9), 2007–2033. https://doi.org/10.1016/j.neubiorev.2012.06.004

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347–367. https://doi.org/10.1177/1094428107308906

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schroeder, P. J. (2014). The effects of age on processing and storage in working memory span tasks and reading comprehension. *Experimental Aging Research*, *40*(3), 308–331. https://doi.org/10.1080/0361073X.2014.896666

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*(1), 4–27. https://doi.org/10.1037/0096-3445.125.1.4

Shamay-Tsoory, S. G., & Mendelsohn, A. (2019). Real-life neuroscience: An ecological approach to brain and behavior research. *Perspectives on Psychological Science, 14*(5), 841–859. https://doi.org/10.1177/1745691619856350

Simon, S. S., Tusch, E. S., Feng, N. C., Håkansson, K., Mohammed, A. H., & Daffner, K. R. (2018). Is Computerized Working Memory Training Effective in Healthy Older Adults? Evidence from a Multi-Site, Randomized Controlled Trial. *Journal of Alzheimer's Disease, 65*(3), 931–949. https://doi.org/10.3233/JAD-180455

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "Brain-Training" programs work? *Psychological Science in the Public Interest, 17*(3), 103–186. https://doi.org/10.1177/1529100616661983

Skevington, S. M., Lotfy, M., & O'Connell, K. A. (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL Group. *Quality of Life Research, 13*(2), 299–310. https://doi.org/10.1023/B:QURE.0000018486.91360.00

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Sunderland, A., Harris, J. E., & Baddeley, A. D. (1983). Do laboratory tests predict everyday memory? A neuropsychological study. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 341–357. https://doi.org/10.1016/s0022-5371(83)90229-3

Sunderland, A., Harris, J. E., & Gleave, J. (1984). Memory failures in everyday life following severe head injury. *Journal of Clinical Neuropsychology*, *6*(2), 127–142. https://doi.org/10.1080/01688638408401204

Tagliabue, C. F., Varesio, G., Assecondi, S., Vescovi, M., & Mazza, V. (2022). Age-related effects on online and offline learning in visuo-spatial working memory. *Aging, Neuropsychology, and Cognition*. https://doi.org/10.1080/13825585.2022.2054926

Tagliabue, C. F., Fornari, C., Varesio, G., & Mazza, V. (2024). Enhancing multiple object analysis skills across early and late adulthood through diverse tasks. *Journal of Cognitive Enhancement*, *8*(3), 220–233. https://doi.org/10.1007/s41465-024-00295-0

Taleb, A., Ismail, A., & Abou-Abbas, L. (2024). Psychometric properties of the Arabic version of the everyday memory questionnaire—Revised (EMQ-R) among the Lebanese population. *The Clinical Neuropsychologist*, *38*(8), 2009–2026. https://doi.org/10.1080/13854046.2024.2343146

Tennstedt, S. L., & Unverzagt, F. W. (2013). The ACTIVE study: Study overview and major findings. *Journal of Aging and Health, 25*(8 0), 3S-20S. https://doi.org/10.1177/0898264313518133

Truong, J., Buschkuehl, M., Smith-Peirce, R. N., Carrillo, A. A., Seitz, A. R., & Jaeggi, S. M. (2022). Change-detection training and its effects on visual processing skills. *Scientific Reports*, *12*(1), 12646. https://doi.org/10.1038/s41598-022-15649-x

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Uittenhove, K., Jeanneret, S., & Vergauwe, E. (2023). From lab-testing to web-testing in cognitive research: Who you test is more important than how you test. *Journal of Cognition, 6*(1), 13. https://doi.org/10.5334/joc.259

van den Noort, M., Haverkort, M., Bosch, P., & Hugdahl, K. (2006). Is there a decline in verbal working memory over age? *Europe's Journal of Psychology*, *2*(2). https://doi.org/10.5964/ejop.v2i2.321

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*(3), 373–403. https://doi.org/10.1016/j.cognition.2014.01.007

Verhaeghen, P., & Basak, C. (2005). Ageing and switching of the focus of attention in working memory: Results from a modified *N*-back task. *The Quarterly Journal of Experimental Psychology Section A, 58*(1), 134–154. https://doi.org/10.1080/02724980443000241

Vermeir, J. F., White, M. J., Johnson, D., Crombez, G., & Van Ryckeghem, D. M. L. (2020). The effects of gamification on computerized cognitive training: Systematic review and meta-analysis. *JMIR Serious Games, 8*(3), e18644. https://doi.org/10.2196/18644

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748. https://doi.org/10.1038/nature02447

Walczyk, J. J., & Taylor, R. W. (1996). How do the efficiencies of reading subcomponents relate to looking back in text? *Journal of Educational Psychology*, *88*(3), 537–545. https://doi.org/10.1037/0022-0663.88.3.537

Wang, G., Zhao, M., Yang, F., Cheng, L. J., & Lau, Y. (2021). Game-based brain training for improving cognitive function in community-dwelling older adults: A systematic review and meta-regression. *Archives of Gerontology and Geriatrics*, *92*, 104260. https://doi.org/10.1016/j.archger.2020.104260

Waris, O., Soveri, A., & Laine, M. (2015). Transfer after working memory updating training. *PLoS One, 10*(9), e0138734. https://doi.org/10.1371/journal.pone.0138734

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., Morris, J. N., Rebok, G. W., Unverzagt, F. W., Stoddard, A. M., Wright, E., ACTIVE Study Group, for the. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA, 296*(23), 2805–2814. https://doi.org/10.1001/jama.296.23.2805

Woods, A. D., Gerasimova, D., Van Dusen, B., Nissen, J., Bainter, S., Uzdavines, A., Davis-Kean, P. E., Halvorson, M., King, K. M., Logan, J. A. R., Xu, M., Vasilev, M. R., Clay, J. M., Moreau, D., Joyal-Desmarais, K., Cruz, R. A., Brown, D. M. Y., Schmidt, K., & Elsherif, M. M. (2023). Best practices for addressing missing data through multiple imputation. *Infant and Child Development, n/a*(n/a), Article e2407. https://doi.org/10.1002/icd.2407

Zelinski, E. M. (2009). Far transfer in cognitive training of older adults. *Restorative Neurology and Neuroscience*, *27*(5), 455–471. https://doi.org/10.3233/RNN-2009-0495

Zelinski, E. M., Gilewski, M. J., & Thompson, L. W. (1980). Do Laboratory Tests Relate to Self-Assessment of Memory Ability in the Young and Old? *New Directions in Memory and Aging (PLE: Memory).* Psychology Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.