The Proportion of Working Memory Items Recoverable from Long-term Memory Remains Fixed Despite Adult Aging

Alicia Forsberg[1,2], Dominic Guitard[2,3], Nathaniel R. Greene[2], Moshe Naveh-Benjamin[2], and Nelson Cowan[2]

[1] Psychology Department, University of Sheffield, England, United Kingdom

[2] Department of Psychological Sciences, University of Missouri, Columbia, MO, USA

[3] Faculte des Sciences de la Sante et des Services Communautaires, Department of Psychology, Université de Moncton, Moncton, New Brunswick, Canada

Author Note

Corresponding author: Alicia Forsberg, Department of Psychological Sciences, University of Missouri-Columbia, 210 McAlester Hall, Columbia, MO, 65211-2500, USA, alicia.forsberg1@gmail.com

Some of these data were presented at virtual conferences in 2021, including a Symposium conducted at the *30th Annual Meeting of the Canadian Society for Brain, Behaviour and Cognitive Science*, the *Virtual Working Memory Symposium*, the *Annual Working Memory Discussion Meeting*, and *Psychonomics*. We thank Bret Glass for assisting in data collection and acknowledge NIH Grant R01 HD-21338, Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD) to Nelson Cowan. Alicia Forsberg developed the study

concept with crucial input from all authors. All authors contributed to the study design. Dominic Guitard programmed the online experiment enabling data collection. Nathaniel Greene performed the Bayesian hierarchical *k* estimation, and Alicia Forsberg performed the remaining data analysis. Alicia Forsberg drafted the manuscript, and all authors provided critical revisions. All authors approved the final version of the manuscript for submission. The methods and all analyses except those labeled 'as exploratory' were pre-registered on the Open Science Framework [https://osf.io/hkezu/?view_only=f71f9e4dfddd4d828238c45db72ce000]. The de-identified data on which the study conclusions are based, study materials, and analytic code needed to reproduce analyses are available at:
https://osf.io/wrt2n/?view_only=e5e3955fc6c241f0a72d757dac65aa4e

Abstract

We explored whether long-term memory (LTM) retrieval is constrained by working memory (WM) limitations, in 80 younger and 80 older adults. Participants performed a WM task with images of unique everyday items, presented at varying set sizes. Subsequently, we tested participants' LTM for items from the WM task and examined the ratio of LTM/WM retention. While older adults' WM and LTM were generally poorer than that of younger adults, their LTM deficit was no greater than what was predicted from their WM performance. The ability to *encode WM information into LTM* appeared immune to age-related cognitive decline.

Keywords: Working Memory; Long-Term Memory; Cognitive Aging

Public Significance Statement: We explored whether older adults' ability to hold information in mind for ongoing cognitive tasks (i.e., their *working memory*) constrains their ability to remember things later on (i.e., their *long-term memory*). For the items that participants could hold in working memory, younger and older adults had comparable rates of long-term recognition, suggesting that avoiding overloading of working memory is a promising approach to helping both younger and older adults remember information.

The Proportion of Working Memory Items Recoverable from Long-term Memory Remains Fixed Despite Adult Aging

Memory is one of the cognitive processes most affected by normal aging (Morcom, 2016; Reuter-Lorenz & Sylvester, 2005), and healthy aging is typically associated with a pronounced decline in both Working Memory (WM) and Long-Term Memory (LTM). WM is a system for temporarily holding mental representations for use in thought and action (Cowan, 2017). The WM system is thought to be limited to 3-4 objects concurrently in young adults (Adam et al., 2017; Cowan, 2001). However, WM appears significantly more limited in older adults (Craik et al., 2010; Chen & Naveh-Benjamin, 2012; Cowan et al., 2006; Gilchrist et al., 2008; Greene et al., 2020; Light & Anderson, 1985; Wingfield et al., 1988). While we seem able to store virtually unlimited information in LTM (Brady et al., 2008), minutes, hours, days, or even years after it was encountered, a large proportion of the information we encounter cannot be retrieved in LTM tests. LTM failures appear especially pronounced in healthy older adults across various materials (e.g., Frieske & Park, 1999; Naveh-Benjamin & Old, 2008; Smith et al., 1990). Despite decades of research on age-related decline in WM and LTM, it is unclear whether those two aging effects are related. The aim of our study was to test this relationship.

The theoretical relationship between WM and LTM is controversial (see Cowan, 2019). Some view the WM and LTM systems as closely related, regarding WM as a temporarily activated, capacity-limited subset of LTM information (e.g., Cowan, 1988, 2019; Morey et al., 2013). Others regard WM and LTM as separate memory systems (e.g., see Shallice & Warrington, 1970; Shimi & Logie, 2018). Early suggestions that WM acts as a bottleneck for LTM encoding (Atkinson & Shiffrin, 1968) have been reinforced in recent work (e.g., Forsberg et al. 2020a; Fukuda & Vogel, 2019, see also Bartsch et al., 2019; Loaiza et al., 2020).

We use the term WM for the maintenance of information needed in the task (in this case, information about visual objects, what some may term 'visual short-term memory'). WM means subtly different things to different researchers, and only some of these definitions have restricted the meaning to situations in which both passive storage and some sort of active processing are required by the task concurrently (Cowan, 2017). Various theoretical frameworks have been constructed in attempts to define how the WM system operates (for an overview, see Logie et al., 2021). For example, the *embedded processes* framework of WM suggests that a central limitation of WM is in the number of items that can be held simultaneously in the focus of attention (Cowan, 1988; 2010; 2019). We use this framework to discuss WM in this paper, but note that we did not study the maintenance of WM bindings – argued by some to characterize the limits of WM capacity (Oberauer, 2005; 2019).

Forsberg et al. (2020a) developed a method to assess the relation between WM and LTM, calculating the ratio of items held in WM that are remembered in a subsequent LTM test (the LTM/WM ratio), under different WM loads. They found that WM capacity limitations constrained the number of items participants recognized in a later LTM test. Moreover, when WM was not overloaded, a higher ratio of items appeared to be encoded in LTM. Based on this evidence of a joint WM and LTM bottleneck in young adults, it is conceivable that older adults' narrower WM capacity bottleneck (e.g., Greene et al., 2020) constrains their LTM performance. Individual differences in WM and speed of processing appear to mediate most age-related variance in LTM (Park et al., 1996, 2002), leading to suggestions that WM plays a crucial role in age-related LTM decline (see Cabeza et al., 2016). Previous research has suggested that when older adults' initial encoding limitations are accounted for, their rate of long-term memory forgetting was similar to that of younger adults (e.g., Kuhlmann et al., 2021; Rybarczyk et al., 1987; Trahan & Larrabee, 1992). Similarly, Puckett and Lawson (1989) observed no age differences in passive forgetting processes over 15 seconds, using word stimuli presented at a rate adjusted to each participants memory span, in a Brown-Peterson task. Although there are previous studies showing that effects of aging of WM can account for effects of aging of LTM using separate WM and LTM materials (e.g., Park et al., 1996), we more directly document the constraints of WM on LTM learning with more direct evidence, examining LTM for the materials studied in the WM task.

Specifically, we used Forsberg et al.'s new method to explore whether the proportion of items held in WM that can later be retrieved from LTM differs with age, distinguishing between three possibilities. First, our participants completed a WM probe-recognition task, and later, their LTM for the WM items was assessed. We used each individual's performance in the WM and LTM tasks to compute the *ratio* of items successfully maintained in the WM phase which were recognized during the later LTM test. This approach differs from previous research in which younger and older adults' initial memory performance was equalized by adjusting the number of memory items studied (e.g., Puckett & Lawson, 1989), or the encoding time (Bartsch et al., 2019). We aim to distinguish between three possible hypotheses.

(1) If older adults' LTM deficits go beyond their WM encoding limitations, they should remember a smaller proportion of their successfully encoded WM items in the subsequent LTM test, compared to younger adults. This might fit with observations of accelerated LTM forgetting in older adults (Baddeley et al., 2014).

(2) If older adults' ability to store items in LTM is intact, but their WM encoding deficits limit the production of LTM representations, younger and older adults may encode a similar proportion of WM items into LTM.

(3) Finally, it is theoretically possible that older adults retain a *higher* proportion of successfully remembered WM items in LTM. This perhaps counterintuitive relationship could occur if older adults exert comparatively more resources to encode fewer items during WM processing (e.g., Cabeza, 2002), in a way that results in deeper LTM encoding (Craik & Lockhart, 1972).

Method

Transparency and Openness

The methods and all analyses except those labeled 'as exploratory' were pre-registered on the Open Science Framework. The de-identified data on which the study conclusions are based, study materials, and analytic code needed to reproduce analyses are available on the Open Science Framework.

Participants

In total, data from 80 younger and 80 older adult participants were included in the experiment. We report some technical pilot data in the Supplement (Section 1). To determine our experimental sample size, we used Bayes Factors design analysis (BFDA; Schönbrodt & Stefan, 2018), taking into consideration detection of group differences or their absence for our most central measure, the LTM/WM ratio. We ran 10000 simulations for a between-subjects design and used a Bayes Factor > 3 as a decision criterion. Two sets of simulations, assuming either no group difference ($d = 0$), or a medium-sized effect ($d = 0.50$), revealed that sequential sampling with a maximum of $N = 80$ per group was appropriate. If there was no effect in these simulations, 88.5% of simulated studies terminated at a H_0 boundary (i.e., correctly finding evidence against an effect), whilst 8.0% of samples were inconclusive (i.e., neither terminating at a H_0 or H_1 boundary), and 3.5% terminated at a H_1 boundary (i.e., a false positive). Conversely, if there was an effect in these simulations, 77.2 % of simulated studies terminated at a H_1 boundary (i.e., correctly finding evidence for an effect), whilst 10.5% of samples were inconclusive, and 12.3% of studies terminated at a H_0 boundary, i.e., false negative). With *N* = 40 per age group, we did find evidence against age differences in the LTM/WM ratio ($BF_{01} > 3$). However, at that point, the evidence regarding age differences in *p*(LTM) was inconclusive ($BF_{01} = 1.13$), suggesting that our data was insufficient to distinguish between there being, or not being, an age effect on *p*(LTM). This prevented us from meaningfully addressing our hypotheses regarding the relationship between older adults' WM and LTM difficulties. We therefore decided to increase our sample to our maximum pre-specified sample size (*N* = 80 participants per age group). Note that while increasing the sample size after analyzing the data is problematic when relying on frequentist hypothesis testing, it is not problematic when using Bayesian statistics (e.g., see Rouder, 2014) and was justified by our optional stopping with maximal *N* rule.

Before reaching *N* = 80 per age group, some participants were excluded and replaced in accordance with our pre-registered exclusion criteria. Specifically, one younger and three older adults were excluded due to performance less than 55% correct at set size 2 in the WM task. The excluded younger adult also failed to meet the second inclusion criteria, as they responded 'old' for more than 90% of new items in the LTM test. A final older adult was excluded and replaced due to a > 10 min break within a memory trial.

We recruited participants via prolific.co (www.prolific.co), for convenient and reliable online data collection found to produce comparable results to in-person laboratory studies (Germine et al., 2012; Peer et al., 2017). This study was approved by the local IRB committee at the University of Missouri and participants were paid $6.76 (USD) upon completion.

Participants reported being native speakers of English; being residents of The United States, Canada, or the United Kingdom; having normal or corrected-to-normal vision; having no cognitive impairment/ dementia and no language-related disorders; and drinking no more than 5-9 alcoholic drinks per week. The mean age of the younger adult group was 21.5 years (*SD*=2.26, range 18–27 years; 51.2% female, 43.8% male, 3.8% other, and 1.3% ‘prefer not to say’). The mean age in the older adult sample was 69.0 years (*SD*=3.23, range 65–77 years; 67.5% female, and 32.5% male). Levels of education in younger and older adults, respectively, were: ‘no formal qualifications’ (1.3, 1.3%), ‘secondary education’ (3.8, 6.3%), ‘high school diploma’ (38.8, 8.8%), ‘technical/community/some college’ (11.3, 25.0%), ‘undergraduate degree’ (36.3, 31.3%), ‘graduate degree’ (6.3, 25.0%), ‘doctorate degree’ (0, 2.5%), ‘don't know/not applicable’ (1.3, 0%), and ‘prefer not to say’ (1.3, 0%). Race reported by younger and older adults, respectively: Asian (13.8, 1.3%), Black or African American (7.5, 2.5%), More than One Race (2.5, 0%), White or European (77.5, 96.3%), Other (6.3, 0%), and ‘prefer not to say’ (2.5, 0%).

Procedure

Participants completed three different tasks: first, a WM probe-recognition task, followed by a 1-minute mathematical distraction task, and, finally, a second probe-recognition memory task, assessing LTM for previously untested items from the WM task. Participants completed all WM trials, and the brief distraction task, before the LTM test. Figure 1 shows example WM (Panel A) and LTM (Panel B) phase trials. The key manipulation was the WM set size (i.e., the number of items presented simultaneously in a circular spatial array; 2, 4, or 6 items). Participants were told that they would complete two memory tests, but were not explicitly told that their memory for the WM items would be tested at a later point. At the very end of the session, we asked them whether they had anticipated a delayed test on these items. We programmed the experiment using PsyToolkit (Stoet, 2010, 2017).

Working Memory (WM) Task

Stimuli were selected from the Microsoft Office ‘Icons’, and consisted of easily recognizable images such as animals, symbols, and furniture. These items were presented in black on a light grey background. Participants studied a total of 288 unique memory items in the WM task, at varying set sizes (2, 4, or 6 items). Each item was presented in one of eight equidistant locations in an imaginary circle around a central fixation cross (+).

Each trial started with a 250 ms central fixation cross, followed by the memory array (presented for 250 ms × the number of items in the array), and a 2000 ms delay, before the probe item and response options were presented. The probe item was drawn randomly from the array on half of the trials (i.e., the *same* as a studied item) and was a new, previously unseen item on the other half (i.e., *different*). Participants responded by clicking on one of the following options presented on the screen along with the probe: ‘I'm sure I saw it’, ‘I think I saw it’, ‘I guess I saw it’, ‘I guess I didn't see it’, ‘I don't think I saw it’, or ‘I'm sure I didn't see it’ (see Figure 1A). The number of trials at the three array sizes was 48, 24, and 16, respectively, resulting in a total of 88 WM test trials, with 96 unique memory items presented at each set size. The order of trials and the selection of items for each trial was randomized for each participant.

Distraction Task

Next, participants completed a 1-minute distraction task, in which they verified mathematical equations of the form $a \times b + c = d$, where a, b, and c were random integers from 1 to 9, and d was equal to $a \times b + c$ or differed from that expression by ±1. Participants responded by clicking ‘Correct’ or ‘Incorrect’ on the screen.

Long-Term Memory (LTM) Task

Finally, participants completed the LTM task. On each trial, participants saw a single probe item and had to say whether they had seen that item in the WM task or not, using the same response scale as in the WM task (see Figure 1B). Items which were probed in the WM task were not probed in the LTM task, to avoid repeated exposure. Similarly, WM lure items were not probed in the LTM task. Each participant responded to a total of 244 items in the LTM task, consisting of 122 new items, 36 set size 2 items, 42 set size 4 items, and 44 set size 6 items.

Results

We use Bayesian statistics, and a nomenclature in which BF_{10} refers to the Bayes Factor for the presence of an effect and BF_{01} refers to absence of an effect, where $BF_{01}=1/BF_{10}$.

Memory Tasks

For each individual, we estimated the proportion of items from a given WM set that was observable in memory at the time of (1) WM testing; *p*(WM), and (2) LTM testing; *p*(LTM). With these measures, we formed a ratio of LTM to WM item production, by set size and age group. Note, though, that the same item was never tested in both WM and LTM.

For these measures, we ignored whether a response was *sure*, *think*, or *guess* and only considered whether the correct half of the response scale was used. The *p*(WM) estimates were obtained by calculating the rate of correct detection of studied items (i.e., the hit rate, *h*) and the rate at which new items were incorrectly identified as old (i.e., the false alarm rate, *f*). The model to estimate *p*(WM) is derived from work by Pashler (1988) and was applied to the present single-probe test situation by Cowan et al. (2013, "reverse-Pashler" formula).[1] We assume that

[1] We found it more natural to redefine *hits* as correct detection of an 'old' or studied item, and *false alarms* as incorrect indications that a novel item was 'old' or studied, differing from Pashler and Cowan et al.

participants respond correctly when the probed item is in WM and otherwise guess that the item is new with a certain rate (*g*). Then, the rate of correct detection of old items, *h*, equals the probability that the probe item is in WM plus the probability that it is not in WM but that a correct "old" guess *g* is given:

$$h=p(\text{WM}) + [1- p(\text{WM})](g)$$

When the item is new there is no match, so performance depends on the guessing rate, and an incorrect response (*f*) is made at the rate, *f*=*g*. Combining these formulas, it can be shown that:

$$p(\text{WM}) = \left(\frac{h-f}{1-f}\right)$$

A comparable formula can also be used to estimate the proportion of items in LTM, in which hits (*hl*) is the rate of correctly detected old items and false alarms (*fl)* the rate of incorrectly responding that a novel item was old:

$$p(\text{LTM}) = \left(\frac{hl-fl}{1-fl}\right)$$

If the false alarm rate exceeds the hit rate, this value will be negative, which makes little sense theoretically and is assumed to involve unfortunate guessing. We had planned two separate approaches to deal with such scenarios. First, we adjusted such values to theoretically plausible values, such that *p*(WM) values equivalent to holding < 1 item in WM were adjusted to equal 1, and *p*(LTM) values < 0 were adjusted to 0, because while a WM capacity of less than one item seems implausible, forgetting all items from a given WM trial in the LTM test seems possible, given the large total number of WM items. In the second approach, participants were completely excluded if they had a negative value at one or more set sizes, either for WM or LTM. Due to space limitations, we report the former, adjusted analyses in the paper, and the latter in the Supplement (Section 2).

Working Memory (WM)

The WM set size manipulation had the expected effect on WM accuracy. Using hierarchical Bayesian logistic regression (Bürkner, 2018) we found credible evidence that memory performance decreased as set size increased, and that younger adults performed better than older adults. See the Supplement (Section 3), for further details about this analysis. An exploratory analysis using '*generalTestBF'* function in the BayesFactor package for R (Morey & Rouder, 2015; R Core Team, 2020) confirmed these patterns for the *p*(WM) measure, as we found 'decisive' evidence for a main effect of Set Size, such that adjusted *p*(WM) was higher for lower-set size items ($BF_{10=}7.64\times10^{53}$); for a main effect of Age, such that younger adults performed better than the older adults ($BF_{10=}2.01\times10^{6}$); and for an Age x Set Size interaction ($BF_{10=}9.95$; see Figure 2, and Table 1). Younger and older adults' performance seems comparable at the lowest set size (likely because younger adults were performing close to ceiling in this condition), but age differences increase as set sizes increases. WM hits rates (i.e., proportion of correctly identified studied items) and WM false alarm rates by age group and set size are presented in Table 1.

Long-Term Memory

We tested whether performance in the LTM task, *p*(LTM), varied as a function of WM set size and age, and whether there was an Age x Set Size interaction. There was 'decisive' evidence for a main effect of Set Size, such that adjusted *p*(LTM) was higher for lower-set size items ($BF_{10=}3.75\times10^{53}$). However, there was inconclusive evidence for the main effect of Age ($BF_{01=}1.05$), and evidence against an Age x Set Size interaction ($BF_{01=}15.25$, see Table 1 for mean values). WM performance at sub-capacity trials (set size 2) was close to ceiling in younger adults (also reflected by the strong evidence for the age group by set size interaction). Younger adults' near-ceiling performance in the set size 2 WM condition might obscure age effects in this condition in the LTM test, if one assumes a maximum WM to LTM transfer ratio in these experimental conditions. Therefore, in an exploratory follow-up analysis, we included only the supra-capacity trials (Set Sizes 4 and 6), and we did observe moderate evidence for age differences in *p*(LTM) (BF_{10}= 3.44). Participants incorrectly responded 'studied' to novel items at the following rates: Younger adults, *M* = .29, *SD* = 0.06; Older adults, *M* = 0.37, *SD* = 0.11. LTM hit rates (i.e., proportion of correctly identified studied items) by age group and set size were: younger adults, Set Size 2: *M* =. 63, *SD* = .16, Set Size 4: *M* = .52, *SD* = .15, Set Size 6: *M* = .44, *SD* = .15; and older adults, Set Size 2: *M* = .65, *SD* = .17, Set Size 4: *M* = .52, *SD* = .18, Set Size 6: *M* = .46, *SD* = .19. Age effects in WM seem more pronounced in the hit rates than for false alarms (see Table 1).

Finally, 25.0% of the younger and 28.8% of the older adults reported having expected that the WM items would be probed again. An exploratory analysis indicated that average *p*(LTM) did not differ between participants who expected LTM test (*M* = 0.33, *SD* = 0.19) and those who did not (*M* = 0.31, *SD* = 0.21; BF_{01} = 5.30, *d* = 0.20).

LTM/WM Ratio: Analysis of the Average Number of Items in WM and LTM

The number of items encoded into WM that could be retrieved on an LTM test (i.e., the LTM/WM ratio) was comparable across age groups. Thus, there was evidence against age differences in the average adjusted LTM/WM ratio (BF_{01}=5.79; younger adult average ratio: *M*=0.40, *SD*=0.16, older adults, *M*=0.41, *SD*=0.17, *d* = .026, see Table 1 for mean ratios by set size). This suggests that, of the items encoded into WM, younger and older adults encoded the items into LTM to similar extents. Figure 2 shows the estimated proportion of items held in WM and LTM, and the LTM/WM ratio, for each WM array size. The ratios for younger and older adults closely coincide with one another. To explore whether this finding was driven by differences in education between the age groups (i.e., older adults having obtained higher levels of education), we replicated the key analysis in a sub-sample of 83 participants who had 'technical/community/some college' (Younger N = 9, Older N = 20) or an 'undergraduate degree' (Younger N = 29, Older N = 25). In this exploratory analysis, we also found evidence against an age difference (BF_{01}=4.00; younger adult average ratio: M=0.39, SD=0.17, older adults, M=0.41, SD=0.17, d = .097). The pattern of results was also found an exploratory analysis of a sub-sample of participants (N = 139, younger N = 62, Older = 77) who were 'White or European' (BF_{01}=4.06; younger adult average ratio: M=0.42, SD=0.13, older adults, M=0.40, SD=0.17, d = .14).

A further question was whether the LTM/WM ratios differed between set sizes and age groups. There was strong evidence that the adjusted LTM/WM ratio differed by Set Size ($BF_{10=}5.15\times10^{10}$), but evidence against an age group difference (BF_{01}=7.00) and against an age group × set size interaction (BF_{01}=8.43). Exploratory comparisons suggested a difference between set sizes 2 and 4 (BF_{10}=1.10×10^{5}, d = 0.53) and between set sizes 2 and 6 (BF_{10}= 1.53×10^{10}, d = 0.74) but inconclusive evidence for difference between set sizes 4 and 6 (BF_{10}=2.32, d = 0.23). Moreover, evidence against an age difference in the LTM/WM ratio was also observed in an exploratory analysis of memory for supra-capacity set sizes (4 and 6 items), BF_{01} = 6.75.

Correlation between Items in WM (k) and LTM

A positive correlation between average WM capacity (k) and the number of items held in LTM is to be expected if WM is the portal to LTM. We indeed observed a positive correlation between age-standardized average *k*-scores and adjusted *p*(LTM) scores, r=.38, BF_{10}=8.50×10^4. For details see the Supplement (Section 5).

Memory confidence ratings

We present confidence ratings for each memory task, set size, and age group in the Supplement (Section 6). The overall rating patterns generally appeared similar in younger and older adults.

Mathematical Distraction Task

On average, younger adult participants attempted 14.4 (*SD*=5.3, range 2–32) problems during this one-minute distraction task, and the average accuracy rate was 77.8% (*SD*=13.1, range 0–95.5% accurate), and older adults completed on average 16.3 (*SD*=3.8, range 8–26) problems, with an average accuracy of 78.6% (*SD*=11.6, range 35.3–92.9% accurate). Exploratory analyses suggested that while older adults completed more math problems than younger adults (BF_{10} = 4.5), there were no age differences in accuracy (BF_{10} = 0.19), indicating that older adults may, on average, have been more engaged in the distraction task or have better arithmetic knowledge.

Discussion

Researchers have attempted to pinpoint key cognitive processes that cause age-related memory deficits. Here, we have isolated the role of WM capacity limitations in age-related LTM decline. Such limitations likely play a key role in many everyday memory situations. We found that when the number of items in a memory array exceeds WM capacity, it not only prevents successful WM retrieval, but also limits subsequent LTM retrieval of items, in both younger and older adults. A similar pattern was observed at an individual level, as participants with higher

WM capacity remembered more items in the LTM task. These findings are in line with recent research with young adults (Forsberg et al., 2020a; Fukuda & Vogel, 2019).

Crucially, we also explored whether the *ratio* of items held in WM that ‘stick’ in LTM differed between age groups. If older adults have additional, distinct LTM deficits (e.g., Baddeley et al., 2014; Park et al., 1996) beyond those shown in WM, they should retain comparatively fewer items during the LTM test. Alternatively, if older adults encode fewer WM items but in a different – perhaps more effortful, deep, manner (Cabeza, 2002) – they might retain a greater proportion of the WM items in the subsequent LTM test. We did not find evidence for either of these possibilities. Instead, we found that while older adults’ WM performance was generally poorer, out of the items they did hold in WM, the ratio of items that could be retrieved at the LTM test was indistinguishable from that of younger adults (see Figure 2). Indeed, in our testing scenario, older adults did not appear to suffer any additional LTM deficits which could not be explained by their WM limitations. Further research should explore whether our results generalize to different LTM paradigms (e.g., using a free recall test, which may produce more marked age-related deficits than our recognition test, e.g., Koen & Yonelinas, 2014; Rhodes et al., 2019; using a longer retention interval, or different memory stimuli). Our memory tests only measured item recognition memory, and findings may not generalize to binding memory (See Bartsch et al., 2019; Oberaurer, 2019). We note similarities with research observing that older adults’ proportional long-term forgetting rate was similar to that of younger adults, when older adults were given fewer items to remember (e.g., Puckett & Lawson, 1989; Rybarczyk et al., 1987; Trahan & Larrabee, 1992).

Moreover, the age-related LTM deficits observed in this study were modest; and only observed at supra-capacity set sizes (4 and 6), leaving it open whether results would replicate in situations where larger age-deficits are found. Interestingly, research using a very similar paradigm suggests that children's WM limitations constrains their LTM performance in a similar way – but again, with less clear evidence at sub-capacity set sizes (Forsberg et al., 2021). Nonetheless, these results have major potential implications, which we discuss next.

First, these results support suggestions that WM limitations contribute directly to age-related LTM deficits (Park et al., 1996), and more specifically, that older adults' LTM deficits might be mitigated by avoiding overwhelming WM during encoding. Theoretically, the role of WM limitations in age-related LTM decline may be closely linked to – rather than compete with – other proposed key causes of LTM decline, such as a decline of inhibitory processes (i.e., Hasher & Zacks, 1988), an associative deficit (i.e., Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008), or declining information processing speed (Salthouse, 1994, 1996). In some WM frameworks, these processes may be considered part of the WM systems, in others, they may be regarded as separate mechanisms supporting performance on WM tasks (see Logie et al., 2021). For instance, older adults' WM limitations may also stem from age-related deficits in processing speed (Brown et al., 2012), and older adults' WM being more easily overloaded may contribute to their associative memory deficits (see Hara & Naveh-Benjamin, 2015).

More broadly, these results are consistent with theories of WM and LTM as linked, closely related systems (Cowan, 2019), which may share the same encoding bottleneck (see Brady et al., 2008). At first glance, our results may seem incompatible with previous work suggesting that successful LTM does not depend on memoranda being held in a WM (or STM) store, but that LTM encoding instead depends on the level of depth and elaboration during encoding (e.g., Craik, 1983). However, the ability to carry out these mental operations depends on sufficient time and the availability of attentional resources. As we consider WM capacity as dependent on focused attention (Cowan, 1988), the constraining factor for holding items in a passive WM – or STM – store in our paradigm, may be the same active, attentional processes, emphasized by Craik (1983). However, our paradigm was designed to test spontaneous WM processing of memory arrays of different set sizes, and thus, while we observed a very clear capacity limitation, the mechanisms driving this limitation are undetermined. Perhaps participants held only a few items in mind and ignored the rest when the number of items exceeded their capacity (e.g., Adam et al., 2017), or perhaps a general resource was distributed across all items, resulting in less precise representations of all items (e.g., Ma et al., 2014). Similarly, we cannot deduce whether participants used strategic approaches of elaboration, verbal rehearsal, or attentional refreshing (Camos et al., 2018). It is possible that such strategies were applied to different extents for arrays of different set sizes, or that strategic approaches differed between age groups (e.g., Craik & Lockhart, 1972; Forsberg et al., 2020b). Moreover, we did not explicitly instruct participants to expect a LTM task, which might increase the use of deep-level strategies, which younger adults may be better positioned to take advantage of (although Fukuda & Vogel, 2019 found no evidence that expecting a LTM test improved young-adult performance, using a similar paradigm). Overall, it is currently an open question whether our pattern of results would replicate using different memory paradigms, which may allow different types of active maintenance or elaborative processes that younger adults might use more successfully than older adults (Craik, 1983). That could produce a higher LTM/WM ratio in younger adults. The importance of our procedure, nevertheless, is to show that there is a baseline condition (i.e., our procedure) that results in equal encoding of WM information into LTM despite any decline in brain function with aging, like the well-known decline in hippocampal function (e.g., Bettio et al., 2017).

Interestingly, Bartsch et al. (2019) found that older adults' WM deficits did not explain their age-related decline in long-term binding memory. While we sought to address the same general question as Bartsch et al. (2019), our paradigms differed in several ways. Bartsch et al. focused on memory for bindings between items – seen by some as crucial to measuring the limits of WM capacity (Oberauer, 2005; 2019), while we tested item recognition. This suggests that long-term binding memory may be especially impaired with age (Old & Naveh-Benjamin, 2008), and that these deficits may be separate from age-related WM deficits. Indeed, binding deficits in WM are not necessarily impaired with age (Allen et al., 2013; Forsberg et al., 2019). Moreover, Bartsch et al. (2019) extended older adults' encoding time (e.g., in Exp. 2, older adults; 1760 ms and, younger adults 710 ms for each word pair) to equate WM performance between groups. In contrast, in our paradigm, participants of both age groups received the same encoding time. The extended encoding time in Bartsch et al. study may have allowed older adults to boost their WM levels to be similar of that of younger adults' but affected their LTM differently than in our paradigm, where encoding time was equally limited for younger and older adults. For example, it might have allowed or encouraged use of a different rehearsal mechanism. However, in our paradigm, older adults may have struggled to perceive – and thus encode – all items. Finally, the use of verbal stimuli (Bartsch et al.) and visual icons (our study) might also have encouraged different rehearsal strategies, or allowed different degrees of elaboration. The discrepancies Bartsch et al. and the current study may reflect important mechanisms driving WM and LTM aging deficits, and future work is needed to pinpoint the underlying mechanisms. Specifically, using our method (equated presentation times per item across age groups) to study binding memory and using the Bartsch et al method (equated performance levels across age groups) to study item memory might help explain understand the observed discrepancies.

As a final note, we want to highlight that although our samples may not be completely representative of the general population (e.g., all but three of the older adults reported being White or European), and the younger and older participant groups were not matched on education level (for example, 6.3% of younger adults had obtained a graduate degree, compared to 25.0% of the older adults), this is the case in many aging studies in which many of the younger participants are still in the process of acquiring higher education, and recruited older adults are mostly white. Furthermore, we have shown that these factors have not mediated our results. Nevertheless, the findings should be replicated in other samples, to ensure their generalizability.

Conclusions

In this study, we examined whether WM limitations constrained LTM and whether items held in WM were encoded in LTM to similar extents in younger and older adults. Experimentally induced WM limitations (i.e., presenting more items than can be held in WM), as well as individual differences in WM capacity, and age-group related capacity differences, all appeared to limit subsequent LTM retrieval. Older adults held fewer items in WM. However, out of the items held in WM, they appeared to encode proportionally as many items into LTM as the younger adults. Thus, when accounting for individual WM capacity limitations, older adults' LTM encoding and recognition processes appeared equal to that of younger adults.

References

Adam, K. C., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, *97*, 79-97. https://doi.org/10.1016/j.cogpsych.2017.07.001

Allen, R. J., Brown, L. A., & Niven, E. (2013). Aging and visual feature binding in working memory. In H. St Clair-Thompson (Ed.), *Working memory: Developmental differences, component processes, and improvement mechanisms* (pp. 83–96). Hauppauge, NY: Nova Science Publishers.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, *2*(4), 89-195. https://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 8, pp. 47–90). New York: Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A., Rawlings, B., & Hayes, A. (2014). Constrained prose recall and the assessment of long-term forgetting: The case of ageing and the Crimes Test. *Memory, 22*(8), 1052-1059. https://doi.org/10.1080/09658211.2013.865753

Bartsch, L. M., Loaiza, V. M., & Oberauer, K. (2019). Does limited working memory capacity underlie age differences in associative long-term memory? *Psychology and Aging, 34*(2), 268. https://doi.org/10.1037/pag0000317

Bettio, L.E.B., Rajendran, L, & Gil-Mohapel, J. (2017). The effects of aging in the hippocampus and cognitive decline. *Neuroscience and Biobehavioral Reviews, 79,* 66-86. doi: 10.1016/j.neubiorev.2017.04.030

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325-14329. https://doi.org/10.1073/pnas.0803390105

Brown, L. A., Brockmole, J. R., Gow, A. J., & Deary, I. J. (2012). Processing speed and visuospatial executive function predict visual working memory ability in older adults. *Experimental Aging Research*, *38*(1), 1-19. https://doi.org/10.1080/0361073X.2012.636722

Bürkner P. C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. http://dx.doi.org/10.18637/jss.v080.i01

Bürkner P. C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*, *17*(1), 85-100. https://doi.org/10.1037//0882-7974.17.1.85

Cabeza, R., Nyberg, L., & Park, D. C. (Eds.). (2016). *Cognitive neuroscience of aging: Linking cognitive and cerebral aging*. Oxford University Press.

Camos, V., Johnson, M. R., Loaiza, V. M., Portrat, S., Souza, A. S., & Vergauwe, E. (2018). What is attentional refreshing in working memory?. *Annals of the New York Academy of Sciences*, *1424*(1), 19-32. https://doi.org/10.1111/nyas.13616

Chen, T., & Naveh-Benjamin, M. (2012). Assessing the associative deficit of older adults in long-term and short-term/working memory. *Psychology and Aging*, *27*, 666–682. http://dx.doi.org/10.1037/a0026943

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163-191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why?. *Current Directions in Psychological Science*, *19*(1), 51-57. https://doi.org/10.1177/0963721409359277

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*, 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin, 145*, 822-847. https://doi.org/10.1037/bul0000199

Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 731-747. https://doi.org/10.1037/a0029687

Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 731-747. https://doi.org/10.1037/a0029687

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Lifespan development of visual working memory: When is feature binding difficult? Developmental Psychology, 42, 1089 –1102. http://dx.doi.org/ 10.1037/0012-1649.42.6.1089

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *302*(1110), 341-359.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I., Luo, L., & Sakuta, Y. (2010). Effects of aging and divided attention on memory for items and their contexts. *Psychology and Aging*, *25*(4), 968-979. https://doi.org/10.1037/a0020276

Forsberg, A., Guitard, D., & Cowan, N. (2020a). Working memory limits severely constrain long-term retention. *Psychonomic Bulletin & Review.* Advance online publication. https://doi.org/10.3758/s13423-020-01847-z

Forsberg, A., Guitard, D., Adams, E. J., Pattanakul, D., & Cowan, N. (2021). Children's Long-Term Retention is Directly Constrained by their Working Memory Capacity Limitations. *Developmental Science*. https://doi.org/10.1111/desc.13164

Forsberg, A., Johnson, W., & Logie, R. H. (2019). Aging and feature-binding in visual working memory: The role of verbal rehearsal. *Psychology and Aging*, *34*(7), 933. https://doi.org/10.1037/pag0000391

Forsberg, A., Johnson, W., & Logie, R. H. (2020b). Cognitive aging and verbal labeling in continuous visual memory. *Memory & Cognition*. 1196–1213. https://doi.org/10.3758/s13421-020-01043-3.

Frieske, D. A., & Park, D. C. (1999). Memory for news in young and old adults. *Psychology and Aging*, *14*(1), 90-98. https://doi.org/10.1037/0882-7974.14.1.90

Fukuda, K., & Vogel, E. K. (2019). Visual short-term memory capacity predicts the "bandwidth" of visual long-term memory encoding. *Memory & Cognition*, *47*(8), 1481-1497. https://doi.org/10.3758/s13421-019-00954-0

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, *19*(5), 847-857. https://doi.org/10.3758/s13423-012-0296-9

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2008). Working memory capacity for spoken sentences decreases with adult ageing: Recall of fewer but not smaller chunks in older adults. Memory, 16, 773–787. http://dx.doi.org/10.1080/09658210802261124

Greene, N. R., Naveh-Benjamin, M., & Cowan, N. (2020). Adult age differences in working memory capacity: Spared central storage but deficits in ability to maximize peripheral storage. *Psychology and Aging*, *35*(6), 866-880. https://doi.org/10.1037/pag0000476

Hara, Y., & Naveh-Benjamin, M. (2015). The role of reduced working memory storage and processing resources in the associative memory deficit of older adults: simulation studies with younger adults. *Aging, Neuropsychology, and Cognition*, *22*(2), 129-154. https://doi.org/10.1080/13825585.2014.889650

Hartshorne, J. K., & Makovski, T. (2019). The effect of working memory maintenance on long-term memory. *Memory & Cognition*, *47*(4), 749-763. https://doi.org/10.3758/s13421-019-00908-6

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The psychology of Learning and Motivation*, *22*, 193-225. https://doi.org/10.1016/S0079-7421(08)60041-9

Koen, J. D., & Yonelinas, A. P. (2014). The effects of healthy aging, amnestic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: A meta-analytic review. *Neuropsychology Review*, *24*(3), 332-354. https://doi.org/10.1007/s11065-014-9266-5

Kruschke, J. K. (2015). *Doing Bayesian data anaylsis. A tutorial with R, JAGS, and Stan.* New York, NY: Academic Press.

Kuhlmann, B.G., Brubaker, M.S., Pfeiffer, T., Naveh-Benjamin, M. (2021). Longer resistance of associative versus item memory to interference-based forgetting, even in older adults. *Journal of Experimental Psychology: Learning Memory and Cognition*, 47(3), 422-438. https://doi.org/10.1037/xlm0000963

Light, L. L., & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. Journal of Gerontology, 40, 737–747. http:// dx.doi.org/10.1093/geronj/40.6.737

Loaiza, V. M., Doherty, C., & Howlett, P. (2020). The long-term consequences of retrieval demands during working memory. *Memory & Cognition*, 1-15. https://doi.org/10.3758/s13421-020-01079-5

Logie, R., Camos, V., & Cowan, N. (Eds.). (2021). *Working memory: The state of the science*. Oxford University Press.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17,* 347-56. https://doi.org/10.1038/nn.3655

Morcom, A. M. (2016). Mind over memory: Cuing the aging brain. *Current Directions in Psychological Science*, *25*(3), 143-150. https://doi.org/10.1177/0963721416645536

Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, *69*(3), 324-348. https://doi.org/10.1016/j.jml.2013.04.004

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs (R package version 0.9.11-1) [Computer software]. Retrieved from http://CRAN.R-project.org/package=BayesFactor

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26,* 1170–1187. https://doi.org/10.1037/0278-7393.24.5.1091

Naveh-Benjamin, M., & Old, S. R. (2008). Aging and memory (humans). In J. H. Byrne, H. Eichenbaum, R. Menzel, H. L. Roediger, & D. Sweatt (Eds.), *Learning and memory: A comprehensive reference* (pp. 787–808). doi:10.1016/B978-012370509-9.00133-9

Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 826-837. https://doi.org/10.1037/0278-7393.29.5.826

Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, *134*(3), 368. https://doi.org/10.1037/0096-3445.134.3.368

Oberauer, K. (2019). Working memory and attention–A conceptual analysis and review. *Journal of Cognition*, *2*(1). http://doi.org/10.5334/joc.58

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., & Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging*, *11*(4), 621-637. https://doi.org/10.1037//0882-7974.11.4.621

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

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153-163. https://doi.org/10.1016/j.jesp.2017.01.006

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125).

Prolific. (n.d.). Prolific.co. www.prolific.co.

Puckett, J. M., & Lawson, W. M. (1989). Absence of adult age differences in forgetting in the Brown-Peterson task. *Acta Psychologica*, *72*(2), 159-175. https://doi.org/10.1016/0001-6918(89)90042-5

R Core Team. (2020). R: A language and environment for statistical computing R version 3.6.3 (2020-02-29) [Computer software]. Vienna, Austria: Author. Retrieved from http://www.R-project.org/

Reuter-Lorenz, P. A., & Sylvester, C. Y. C. (2005). The cognitive neuroscience of working memory and aging. In R. Cabeza, L. Nyberg, & D.Park (Eds.), Cognitive neuroscience of aging: Linking cognitive and cerebral aging (pp. 186–217). Oxford, England: Oxford University Press.

Rhodes, S., Cowan, N., Hardman, K. O., & Logie, R. H. (2018). Informed guessing in change detection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1023-1035. https://doi.org/10.1037/xlm0000495

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin & Review*, *26*(5), 1529-1547. https://doi.org/10.3758/s13423-019-01649-y

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301-308. https://doi.org/10.3758/s13423-014-0595-4

Rouder, J. N., Morey, R. D., Speckman, P. L., & Pratte, M. S. (2007). Detecting chance: A solution to the null sensitivity problem in subliminal priming. *Psychonomic Bulletin and Review*, *14*, 597-605. https://doi.org/10.3758/BF03196808

Rybarczyk, B. D., Hart, R. P., & Harkins, S. W. (1987). Age and forgetting rate with pictorial stimuli. *Psychology and Aging*, *2*(4), 404. https://doi.org/10.1037/0882-7974.2.4.404

Salthouse, T. A. (1994). The nature of the influence of speed on adult age differences in cognition. *Developmental Psychology*, *30*(2), 240-259. https://doi.org/10.1037/0012-1649.30.2.240

Schönbrodt, F. D. & Stefan, A. M. (2018). BFDA: An R package for Bayes factor design analysis (version 0.3). Retrieved from https://github.com/nicebread/BFDA

Shallice, T., & Warrington, E.K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22*, 261-273. https://doi.org/10.1080/00335557043000203

Shimi, A., & Logie, R. H. (2019). Feature binding in short-term memory and long-term learning. *Quarterly Journal of Experimental Psychology*, *72*(6), 1387-1400. https://doi.org/10.1177/1747021818807718

Smith, A. D., Park, D. C., Cherry, K., & Berkovsky, K. (1990). Age differences in memory for concrete and abstract pictures. *Journal of Gerontology*, *45*(5), P205-P210. https://doi.org/10.1093/geronj/45.5.p205

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

Su., Y.-S., & Yajima, M. (2015). R2jags: Using R to Run ‘JAGS’ [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=R2jags (R package version 0.5-7).

Trahan, D. E., & Larrabeereal, G. J. (1992). Effect of normal aging on rate of forgetting. *Neuropsychology*, *6*(2), 115–122. https://doi.org/10.1037/0894-4105.6.2.115

Wingfield, A., Stine, E. A., Lahar, C. J., & Aberdeen, J. S. (1988). Does the capacity of working memory change with age? Experimental Aging Research, 14, 103–107. http://dx.doi.org/10.1080/03610738808259731

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition (2nd ed.)*. (pp. 293-357). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Table 1. *Averages for transformed data by Set Size (adjusted; N = 80 per age group)*

Set Size	*p*(WM)	*p*(LTM)	LTM/WM ratio	WM Accuracy (Proportion correct)	WM hit rates	WM false alarm rates
Younger Adults						
2 items	.95 (.06)	.48 (.22)	0.50 (0.22)	.97 (.04)	.96 (.05)	.02 (.06)
4 items	.83 (.16)	.31 (.17)	0.38 (0.21)	.89 (.09)	.85 (.14)	.07 (.10)
6 items	.70 (.20)	.21 (.13)	0.33 (0.25)	.81 (.11)	.73 (.18)	.11 (.13)
Older Adults						
2 items	.89 (.11)	.46 (.20)	0.51 (0.21)	.93 (.06)	.90 (.11)	.03 (.05)
4 items	.67 (.20)	.25 (.15)	0.38 (0.24)	.81 (.10)	.69 (.18)	.07 (.10)
6 items	.54 (.23)	.17 (.14)	0.33 (0.26)	.74 (.12)	.59 (.22)	.11 (.13)

Note. Values in Parenthesis represent Standard Deviations. Values equivalent to WM $k < 1$ were replaced with the corresponding p(WM) for $k = 1$, and values of $p(\text{LTM}) \times \text{Set Size} \leq 0$ were replaced with zero. WM hit rates = the proportion of correctly identified studied items.

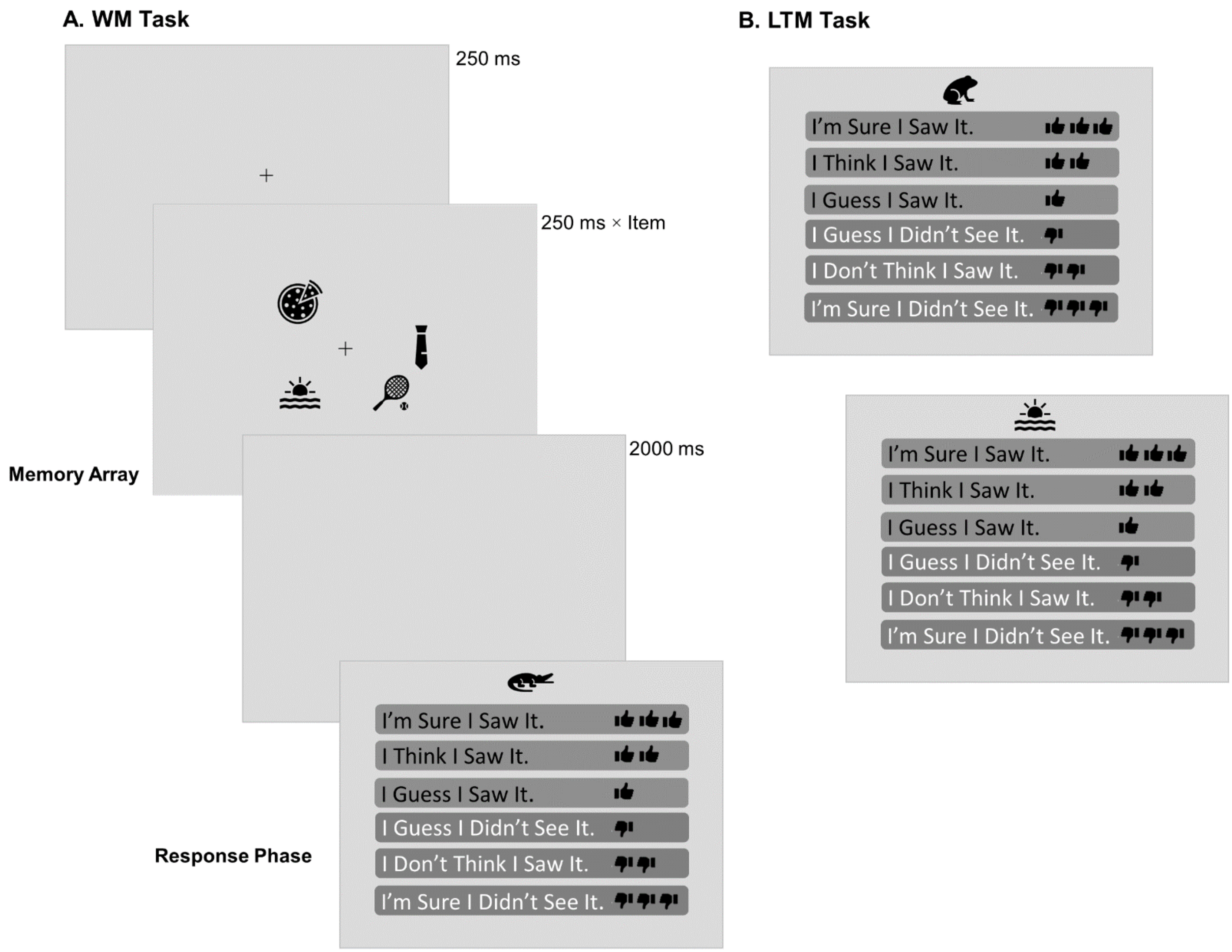

Figure 1. Outline of some typical trials. **Panel A**, Working Memory (WM) Task trial. **Panel B**, two trials in the Long-Term Memory (LTM) Task. The memory array set size in the WM task varied between 2, 4, or 6 items, and the presentation time was adjusted to be 250 ms per item. During the WM response phase, participants indicated whether they had seen the probe item in the Memory Array or not, and also selected their level of confidence, by a mouse-click on the relevant option. In the LTM task, participants indicated whether they had previously seen the probe items in the WM task. This figure is presented in grey-scale. In the real experiment, the response scale was green for the three response options indicating having seen the object, and red for the options indicating not having seen the objects.

Figure 2. Memory accuracy by WM set size. **Panel A**, *p*(WM) (i.e., the estimated probability that items were held in WM, calculated using the "reverse-Pashler" formula, Cowan et al., 2013); **Panel B**, *p*(LTM) (i.e., the estimated probability that items were held in LTM); **Panel C**, the p(LTM)/p(WM) ratio. Triangles and the dashed line show the young adults' performance, the circles show the older adults' performance. Lighter circles and triangles show individual subjects estimates (the points are jittered slightly in the figure to avoid overlap). Error bars represent 95% confidence intervals.