This is a repository copy of *Semantic relatedness corrects the age-related binding deficit in working memory and episodic memory*.

Article:

Loaiza, V.M. orcid.org/0000-0002-5000-7089 and Srokova, S. (2020) Semantic relatedness corrects the age-related binding deficit in working memory and episodic memory. The Journals of Gerontology: Series B, 75 (9). pp. 1841-1849. ISSN 1079-5014

https://doi.org/10.1093/geronb/gbz055

This is a pre-copyedited, author-produced version of an article accepted for publication in The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences following peer review. The version of record Loaiza, V.M. and Srokova, S. (2020) Semantic relatedness corrects the age-related binding deficit in working memory and episodic memory. The Journals of Gerontology: Series B, 75 (9). pp. 1841-1849. is available online at: https://doi.org/10.1093/geronb/gbz055.

Semantic Relatedness Corrects the Age-Related Binding Deficit in Working Memory and Episodic Memory

Vanessa M. Loaiza, Ph.D.

Sabina Srokova, B.Sc.

Manuscript accepted for publication at *Journals of Gerontology, Series B: Psychological Sciences* on 27 April 2019.

Word count: 4,999 (title page, abstract, and text)

Author Note

Vanessa M. Loaiza, Department of Psychology, University of Essex. Sabina Srokova is now at the University of Texas, Dallas.

Correspondence concerning this article should be addressed to Vanessa M. Loaiza, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: v.loaiza@essex.ac.uk

This research was funded by an internal impact grant from the Research Enterprise Office at the University of Essex awarded to Vanessa M. Loaiza. The authors acknowledge Enrique Arcilla, Maria Sanz Taberner, and Jia Ching See for their assistance with data collection.

The pre-registration, materials, data, and analysis scripts for the study are available at the Open Science Framework: https://osf.io/keu4d/

Abstract

Word count: 185/200

Objectives: It is well-known that age differentially impacts aspects of long-term episodic memory (EM): Whereas a binding deficit indicates that older adults are less capable than younger adults to encode or retrieve associations between information (e.g., the pairing between two memoranda, such as *lock – race*), item memory is relatively intact (e.g., recognizing *lock* without its original pairing).

Method: We tested whether this deficit could be corrected by facilitating establishment of the bindings in working memory (WM) through adapting the semantic relatedness of studied pairs according to participants' ongoing performance (Experiments 1 and 2). We also examined whether this was evident for the long-term retention of pairs that were not tested in WM (Experiment 2).

Results: The results revealed matched binding and item memory in WM and EM between age groups. Most importantly, older adults required increased semantic strength between word pairs to achieve similar performance to that of younger adults, regardless of whether pairs were immediately tested during the WM task.

Discussion: These findings indicate that relying on their superior semantic memory can correct the commonly exhibited profound deficit in binding memory in older age.

Keywords: associative memory, episodic memory, semantic memory, working memory

Memory ability is often thought to decline as adults grow older. However, this decline is not absolute: Whereas age-related deficits in the retention of episodes and events, or *long-term episodic memory* (EM), are well-known (McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010; Verhaeghen & Salthouse, 1997), the accumulated knowledge of facts and concepts, or *long-term semantic memory* (SM), excels with age (Park et al., 2002; Verhaeghen, 2003). Within the domain of EM, the retention of specific, contextual details of an event, or *binding memory*, shows age-related impairments, yet memory in the absence of such details, or *item memory*, is relatively preserved (McCabe, Roediger, McDaniel, & Balota, 2009; Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995). Substantial work thus contradicts the conventional view of unequivocal age-related memory decline. Nevertheless, understanding the nature of the memory difficulties in older age is a high priority, not just for clarifying theoretical conceptions more broadly, but to also address how can memory be improved in older age.

According to the *associative deficit hypothesis*, older adults are less able to build and retrieve bindings between components into a cohesive unit, whereas recognition of the individual items themselves is less impaired (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). For example, Naveh-Benjamin (2000) instructed participants to study word pairs (e.g., *lock – race*, *zoo – time*, *pen – truck*) for two types of tests: Either an item recognition test, wherein individual items (e.g., *zoo, truck*) were tested, or a pair recognition test, wherein intact (e.g., *lock – race*), rearranged (e.g., *pen – time*), or entirely new pairs (e.g., *visitor – shape*) were tested. Both types of tests were then administered to all the participants. Older adults performed worse than younger adults overall, but the deficit was larger for the pair recognition test than the item recognition test. Furthermore, the forewarning of the type of test only benefitted younger adults during the pair recognition test, suggesting that older adults cannot capitalize on strategic resources during encoding.

Such findings indicate that older adults have fewer attentional resources in working memory (WM) compared to their younger counterparts, with downstream consequences for EM. Indeed, WM deficits are evident in older age (Bopp & Verhaeghen, 2005) and account for substantial age-related variability in EM (McCabe et al., 2010; Park et al., 2002; Verhaeghen & Salthouse, 1997), but thus far the causative role remains unclear. Among the competing accounts (see Loaiza & Oberauer, 2016 for a review), one view emphasizes the role of WM to establish, maintain, and manipulate short-term content-context bindings (Oberauer, 2005). Thus, reduced WM capacity to build and maintain bindings may consequently impair their retention in EM.

Often researchers have varied attentional demand during encoding to investigate whether distracting attention reduces WM resources that contribute to formation of bindings (Hara & Naveh-Benjamin, 2015; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003). However, a disadvantage of this approach is that age differences often already exist in the baseline full attention condition, and thus an ordinal interaction may occur due larger age differences under dived attention. Such ordinal interactions are common but difficult to interpret (Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2012). An ordinal interaction may occur simply because of the scale of measurement: There may be greater sensitivity at one end of the scale than the other, thereby causing an effect (e.g., an age difference) to appear larger or smaller than it actually is under different conditions (e.g., full versus divided attention). Furthermore, the binding deficit is inconsistently simulated in studies of divided attention in younger adults (Craik, Luo, & Sakuta, 2010; Kilb & Naveh-Benjamin, 2007).

Rather than impairing performance, the reverse approach of improving older adults' performance to that of younger adults would avoid these issues and suggest theory-based methods to improve memory in older age. Bartsch, Loaiza, and Oberauer (2019) recently tested a novel adaptive relational recognition paradigm that equated binding in WM between age groups to determine whether the corresponding deficit in EM is reduced. This paradigm presented several unrelated word pairs (e.g., *lock – race*, *zoo – time*), and afterward one of the items (e.g., *lock*) cued recall of its pair from three options: The correct item (e.g., *race*), an incorrect new item (e.g., *visitor*), or an incorrect lure item that was presented in the trial but not with that item (e.g., *time*). If participants' binding performance fell below 67%, the presentation rate of the pairs for the subsequent trial became slower, whereas exceeding the criterion meant that the subsequent trial was faster. As such, binding in WM was relatively matched between groups, although older adults needed substantially slower presentation rates to than younger adults to achieve this. Furthermore, the age deficit in retention of the bindings in EM was greatly reduced, suggesting that a slower presentation rate mitigated the binding deficit.

This promising method of tailoring the task on an individual basis is particularly relevant given the growing area of research concerning whether older adults' aforementioned greater SM may buffer their EM deficits (see Umanath & Marsh, 2014 for review). Although it is well-known that prior knowledge and elaboration improve memory performance overall, researchers have investigated whether older adults disproportionately benefit relative to younger adults (Badham, Hay, Foxon, Kaur, & Maylor, 2015; Badham & Maylor, 2015; Castel, 2005; Mohanty, Naveh-Benjamin, & Ratneshwar, 2016; Naveh-Benjamin, 2000). However, as mentioned previously, much of this literature relies on ordinal interactions to assess whether semantic meaningfulness is especially important to older adults. Thus, unambiguous evidence is still necessary to show that older adults are particularly sensitive to factors that draw upon their superior SM to buffer the binding deficit in EM.

To this end, we modified Bartsch and colleagues' adaptive relational recognition task to equate binding ability in WM between age groups by adjusting the semantic relatedness of the studied word pairs (e.g., presenting *lock – vault* instead of *lock – race*) based on ongoing task performance. We predicted that older adults should rely on semantic meaningfulness (i.e., increased number of related pairs per trial) to a greater extent than younger adults to achieve similar binding memory in WM. If creation of bindings in WM is important to their retention in EM, then the age-equivalent binding memory should hold over the long-term.

Experiment 1

Method

Participants. We recruited sixty-one participants (30 younger and 31 older adults) for the study (see Table 1). Participants reported generally good health with no medical history of memory or cognitive impairment. One older adult was excluded for failing to pass the mini mental status examination (MMSE; Folstein, Folstein, & McHugh, 1975). Participants were recruited from the local community and were native or fluent English speakers. Participants provided informed consent and were fully debriefed and compensated £6 for their participation. The ethics committee of the University of Essex approved the ethics application for both experiments.

Materials and Procedure. The experiment was programmed in Matlab using the Psychtoolbox extensions (Brainard, 1997; Kleiner et al., 2007) and closely followed the procedure of Bartsch and colleagues (2019). The older adults first completed the MMSE. The participants were presented with an adaptive relational recognition task consisting of five blocks of five WM trials (i.e., 25 trials total), with each block followed by a distraction task and EM test. The memoranda were drawn from a pool of 11,963 pre-selected pairs of nouns (letters: $M = 5.49$, $SD = 1.81$, range = 2 – 14; syllables: $M = 1.64$, $SD = 0.79$, range = 1 – 5; log HAL frequency: $M =$

9.52, *SD* = 1.68, range = 3.09 – 14.35), representing a subset of normed cue-target pairs constructed by Nelson, McEvoy, and Schreiber (2004). For related pairs, the backward associative strength was 0 and the forward associative strength was between .02 – .04.

The experiment was programmed so that combinations of five words were randomly drawn from the set for each participant: the cue, a related potential target, an unrelated potential target, and two new words to serve as the negative response options during the WM and EM phases of the experiment, respectively. The stimuli were all unique (i.e., individual words or pairings were never repeated in the experiment). The unrelated pairs were randomly arranged with the additional constraint that they were not arranged into pairs that were in fact inadvertently related. The relatedness of the pairs within each trial was also unique (e.g., if *feeling – bad* was a possible related pair in the trial, then *shame – abuse* would not appear in the same trial because *shame – bad* is a related pair). The unrelated and new items were also selected with this constraint, such that they were not semantically related to any of the other items of the pairs within the same trial.

Participants were tested individually in quiet booths. An experimenter was present during the instructions and practice trials and monitored several participants from outside their booths thereafter. The participants first completed two practice trials. A fixation point was presented at the center of the screen for 1000 ms, followed by successive presentation of five word pairs (e.g., *lock – race*, *zoo – time*) presented for 1000 ms each. The pairs were presented from the top to the bottom of the screen, and the participants were instructed to silently read and remember them. Immediately after, all the pairs were tested in random order (unspeeded responding). For each pair, the probe word (i.e., the cue) was presented at the center of the screen (e.g., *race*), and participants indicated the correct option (i.e., the target) from a random arrangement of three possible responses below it: positive (i.e., the correct word, e.g. *lock*), negative (i.e., a never-presented item, e.g. *sign*), and intrusion (i.e., a lure presented during the trial, but not in that pair, e.g., *zoo*).

Following Bartsch and colleagues, an adaptive algorithm was active across all 25 trials of the task and continuously monitored ongoing accuracy over a moving window of the previous 10 trials, such that after each trial, performance was averaged across up to the previous 10 trials. The experiment began with all unrelated pairs, and the adaptive algorithm adjusted the number of semantically related pairs presented during the subsequent trial by comparing the ongoing accuracy to a criterion. Based on a pilot study, the criteria were 60% and 67% correct responses (i.e., positive response options) for younger and older adults, respectively, with the aim of reaching 67% accuracy in both age groups.[1] Thus, if the participants' average accuracy fell under their age group's criterion, the subsequent trial introduced a related word pair. In contrast, an unrelated pair replaced a related word pair when the average performance exceeded the criterion, unless there were no related pairs, in which case subsequent trials continued to comprise only unrelated pairs.

After each block, a distraction phase of simple arithmetic problems was completed for 1 min, followed by the EM test of the word pairs from the previous block. Once again, each probe word was presented in random order with three response options (positive, negative, intrusion) that were randomly displayed below it. The same cue word and correct target from the WM trials were presented again, but the intrusion and negative options were a novel combination than what was presented during the WM trials. That is, the intrusion option represented a different item that was presented during the WM test from the same trial, and the negative option was an entirely new

---

[1] This difference in criteria was merely a pragmatic one in line with our pre-registered aim to achieve equivalent WM performance between age groups. Furthermore, the approach was analogous to that of Bartsch and colleagues, who also showed that the adaptive algorithm tended to overestimate younger adults' performance when implemented identically to the older adults.

word to the experiment. After completing all blocks, the participants completed a computerized demographic questionnaire and vocabulary test (Shipley, Gruber, Martin, & Klein, 2009).

Design. The independent variable of the experiment was age group (younger versus older adults). There were three observed measures: The first two were measures of recall accuracy during both the WM and EM tasks. The strict score represented correctly selecting the positive response option, whereas the lenient score considered both the positive and intrusion response options as correct. As detailed previously, the strict score was used to determine the introduction or removal of a semantically related pair to the subsequent trial based on ongoing performance. Although we report these scores in line with our pre-registration, we focused our interpretation on the estimated binding and item memory parameters that are explained in the next section. Note that the strict and lenient scores are different than how the binding and item parameter estimates were derived, as we explain further in the next section. Finally, the third measure concerned the proportion of semantically related word pairs that were presented during the trials.

Data Analysis. According to our pre-registration, we compared younger and older adults using Bayesian *t*-tests (Morey & Rouder, 2015) and Bayesian Estimation Software (BEST; Kruschke, 2013) in R (R core team, 2017) to examine whether there is similar binding memory between age groups (i.e., a null effect of age). Bayesian statistics compare the likelihood of the data given a particular model, such as a model that assumes a null effect of age group, to that of another model, such as the alternative model that assumes a main effect of age group. The ratio of the likelihood of these models is the Bayes Factor (BF) that quantifies the evidence for accepting one model over another model (e.g., the ratio of the alternative model to the null model, BF_{10}). BFs between 1 and 3 are considered relatively ambiguous, whereas BFs greater than 10 and 100 are considered as strong and decisive evidence, respectively. Alternatively, one could examine the size of the age effect and the probability of values that fall within a region of practical equivalence (ROPE; effect sizes within the range of -0.1 and 0.1). We report both BFs and BEST for the observed dependent measures.

Additionally, we fit the data to separate hierarchical Bayesian multinomial processing tree (MPT) models for each age group using the TreeBUGS package (Heck, Arnold, & Arnold, 2017), following the analysis procedures of Bartsch and colleagues. MPT models estimate the probability of underlying latent cognitive parameters through raw categorical data. Rather than observed average accuracy (e.g., where only positive responses are considered correct in the case of the "strict score"), the MPT approach uses the observed frequencies for each response category (i.e., the positive, negative, and intrusion response options). This allowed precise measurement of binding and item memory while also explicitly accounting for participant heterogeneity and guessing. We used the same MPT model as Bartsch and colleagues (see Figure 1): The model first defines whether or not participants have accurate binding memory for the tested pair (with probability $P_{Binding}$). In the absence of binding memory ($1 - P_{Binding}$), the participants may still have accurate item memory (P_{Item}) and are therefore able to distinguish the two items that were presented in the trial from the negative option and guess between them. This would lead to either a correct positive response ($g_B = 0.5$) or to an incorrect intrusion response ($1 - g_B$). In the absence of item memory ($1 - P_{Item}$), participants would guess among all three options with equal probability ($g_i = 0.33$ where i refers to three possible guessed options). For both analysis of the raw data and the MPT models we used default priors of the respective analysis packages and computed the credibility intervals and mean parameters from the posterior samples.

Results and Discussion

In line with Bartsch and colleagues (2019), performance on the first block (i.e., five trials) was excluded from analysis. Estimated mean performance (and corresponding 95% highest density intervals, HDIs) are presented in Table 2. We first examined the success of the algorithm to match performance between age groups in terms of the WM strict score. There was ambiguous evidence for an age-related benefit, such that older adults slightly outperformed younger adults. Furthermore, mean performance was near the criterion of approximately 67%, showing that the algorithm was successful for matching age groups on the strict score. Importantly, there was decisive evidence for an age difference in the proportion of semantically related pairs, such that older adults required more related pairs than younger adults to achieve a similar strict score.

Given that the observed strict and lenient scores are not pure estimates of binding and item memory, respectively, we followed the analysis protocol of Bartsch and colleagues to also examine the latent cognitive states of binding and item memory as derived from the hierarchical MPT model described previously. The model converged ($\hat{R}$s < 1.01) and provided a good fit of the data for both age groups, with observed and predicted values showing little deviance and the posterior predictive *p* (PPP) values were large (PPPs > .39). The posterior distribution of the parameter estimates for each age group are shown in Figure 2. The parameter estimates for all the parameters were overlapping, with older adults showing a slight advantage in binding over younger adults. Examining the mean differences and 95% credibility intervals of the age comparisons confirmed the null age difference in each parameter: The difference between the posteriors centered on 0 for binding memory in WM (mean = -0.05, CI = [-0.11, 0.01]) and EM (mean = -0.05, CI = [-0.13, 0.03]), and item memory in WM (mean = -0.03, CI = [-0.19, 0.14]) and EM (mean = 0.08, CI = [-0.11, 0.25]). These results support our hypothesis that binding memory can be matched between age groups by adapting the semantic relatedness of the presented bindings, with older adults requiring more pairs to be related than younger adults.

These results support the notion that semantic relatedness can be adapted to support the encoding and long-term retention of bindings in older age. However, it may be the case that the null age effect persisted in EM as a result of facilitated retrieval rather than bolstered creation of bindings during encoding. That is, the adapted semantic relatedness may allow participants to simply recall which target appeared with the cue during immediate recall in the WM task.

In order to address this possibility, we conducted a second pre-registered experiment with a very similar design to the first, with the exception that only a random half of the presented pairs in each trial were tested, and all the pairs were later tested in EM. We expected to observe a testing effect, such that participants should show greater EM for the pairs tested in WM relative to the untested pairs. We also expected to replicate the results of Experiment 1, such that there should be a null age effect in binding memory for pairs tested in WM that is achieved via a relatively greater proportion of presented related pairs for older than younger adults. The most important result concerned the effect of age on untested pairs: If participants simply remember what they had retrieved during WM, then the null age effect in EM binding memory should be exclusive to pairs that had been tested in WM, whereas an age deficit should persist for untested pairs. Conversely, if the benefit of semantic relatedness concerns the creation of bindings in WM, then there should be a null age effect in EM binding memory regardless of whether the pairs were tested during WM.

Experiment 2

Method

Participants. We recruited 31 younger and 33 older adults who did not participate in but were otherwise similar to the participants of Experiment 1 (see Table 1). The data of one older adult were excluded for failing to pass the MMSE. Participants provided informed consent and were fully debriefed and compensated £7.50 for their participation at the end of the experiment.

Materials and Procedure. The materials and procedure were very similar to Experiment 1, with the following exceptions: Participants were immediately tested on a random three of six presented word pairs per trial, and all the pairs thereafter during the EM test. Thus, there were 30 total presented pairs during each of the five blocks, with half untested during WM. The intrusion options were arranged so that they were exclusive to the word pairs that were tested versus untested in WM. For example, if the word pairs *destiny – mirror, country – diamond,* and *nylon – paper* were presented and immediately tested, then the intrusion options could be, respectively, *diamond, paper, mirror* in the WM test and *paper, mirror, diamond* in the EM test, with the other three untested word pairs from that trial serving as their own intrusions during the EM test. Furthermore, the adaptive algorithm operated by aiming to evenly present any change in the relatedness between the tested and untested pairings. If a related pair was introduced or removed in the next trial, the algorithm would first check the count of related tested and untested pairs thus far. For example, if recall fell below the criterion and no related pairs had yet been introduced, then the algorithm would randomly determine whether the to-be-introduced related pair in the next trial would be tested or untested, but would assign the to-be-introduced related pair to be tested if two related pairs were untested and one was tested in the last trial. This allowed the experiment to achieve approximately even presentation of related pairs between those that were tested versus untested in WM. The rest of the experiment and analyses were the same as Experiment 1.

Results and Discussion

As in Experiment 1, we first examined the success of the algorithm to match the criterion WM strict score between age groups and the proportion of presented pairs that were related (see

Table 2). Once again, the WM strict score was similar between age groups, and older adults required more semantically related pairs than younger adults. Somewhat unexpectedly, the age difference in the related pairs was not as overwhelming as in Experiment 1, although it was still credible. A follow-up exploratory analysis suggested the difference between experiments was not credible, with ambiguous evidence for a null effect of experiment in both age groups (Younger adults, $BF_{10} = 0.50$; older adults, $BF_{10} = 0.48$). In all, the pattern of Experiment 1 was replicated, such that the algorithm successfully matched age groups on the WM strict score by administering more related word pairs to older than younger adults.

We next applied the MPT model to the data to ascertain parameter estimates for binding and item memory (see Figure 3). The model successfully converged ($\hat{R}$s < 1.02), and the fit was good for both age groups (PPPs > .14). Consistent with Experiment 1, binding memory in WM was similar between age groups (mean age difference in posteriors = -0.02, CI = [-0.08, 0.04]). Furthermore, an advantage for tested pairs over untested pairs was evident in binding memory in EM for younger (mean = 0.18, CI = [0.11, 0.26]) and older (mean = 0.10, CI = [0.03, 0.16]) adults. This was not the case for item memory (younger: mean = 0.08, CI = [-0.08, 0.24]; older: mean = 0.09, CI = [-0.06, 0.24]), although low item memory overall may have obfuscated any testing effect. Regarding the age effects in EM, we replicated the null age difference in binding memory (mean = -0.03, CI = [-0.12, 0.07]) and item memory (mean = -0.01, CI = [-0.20, 0.18]) for pairs tested in WM. In summary so far, the results of Experiment 2 replicated those of Experiment 1.

The critical analysis concerned parameter estimates, particularly binding memory, for untested pairs. Binding memory for untested pairs was credibly greater in older than younger adults (mean = -0.11, CI = [-0.20, -0.03]), whereas there was no age difference in item memory for untested pairs (mean = 0, CI = [-0.13, 0.13]). Although not predicted, the age benefit in binding memory may have been anticipated in hindsight given that older adults studied more related pairs than younger adults, thereby yielding less forgetting of those pairs in older adults. Further exploratory analysis applied the MPT model to the data of only the related pairs of the experiment.[2] The model converged ($\hat{R}$s < 1.04) and fit well to the data of both age groups (PPPs > .13). Most importantly, the age differences centered on 0 for binding memory in WM (mean = -0.09, CI = [-0.21, 0.03]), binding memory in EM for tested pairs (mean = -0.04, CI = [-0.18, 0.09]), and binding memory in EM for untested pairs (mean = -0.10, CI = [-0.24, 0.09]). That binding memory was the same or better for untested pairs suggests that the benefit of semantic relatedness was not merely due to facilitation of retrieving the cue-target pairs during WM. Thus, these results support the notion that adapted semantic relatedness benefits older adults' long-term retention by bolstering the encoding and maintenance of bindings in WM.

General Discussion

The present study examined the importance of establishing and maintaining bindings in WM for the age-related associative deficit in EM. Instead of the typical approach of varying attentional demand (Hara & Naveh-Benjamin, 2015; Naveh-Benjamin et al., 2003), we focused on enhancing older adults' memory performance to equal that of younger adults by adapting the semantic relatedness of the memoranda according to their ongoing performance. Given their relatively superior SM (Park et al., 2002; Verhaeghen, 2003), it was expected that older adults would disproportionately benefit from semantic support relative to younger adults. As predicted, older adults required a greater number of semantically-related word pairs to achieve similar binding memory in WM as younger adults, and importantly, the age-equivalent binding memory

---

[2] It should be noted that this resulted in a loss of data, particularly of the eight younger adults who never required related pairs.

was retained when the pairs were later retrieved again from EM. Thus, the age-related associative deficit in EM can be ameliorated by capitalizing on the extensive and densely networked SM that older adults have acquired over their lifetimes.

These findings are consistent with prior studies indicating that older adults can rely on their superior SM to overcome age-related memory deficits and may differentially benefit from semantic support relative to younger adults (Badham et al., 2015; Badham & Maylor, 2015; Castel, 2005; Mohanty et al., 2016). The results also cohere with the first use of this adapted relational recognition task (Bartsch et al., 2019). Specifically, Bartsch and colleagues showed that correcting the WM binding deficit by adapting the presentation rate of presented word pairs greatly reduced the corresponding age-related binding deficit in EM. Furthermore, although varying the number of presented pairs (i.e., set size) strongly affected WM binding, set size had no effect on retention of the bindings in EM for either age group. These results suggested that age-related deficits in WM do not cause the binding deficit in EM, rather, binding deficits may generally reflect a common cause of inefficient encoding. The authors suggested that slower processing speed in older age (Salthouse, 1996) was a possible explanation, but others, such as poor discrimination of bindings or impaired consolidation of stable memory representations, could not be ruled out.

At first glance, the current research appears inconsistent with the processing-speed theory because we observed a similar pattern of results without any manipulation of the presentation rate. Furthermore, the equated binding memory demonstrated in WM held for EM, whereas Bartsch and colleagues showed that an age deficit in EM still persisted using the presentation rate adaptation. Rather than a cause per se, slower presentation rates more likely reflect the underlying operations required to encode and retain bindings regardless of whether they are retrieved from WM or EM. On the other hand, the relatedness of the pairs may have facilitated their binding in the limited time available during presentation. It is difficult to disambiguate these possibilities. Generally, the current results cohere with a common cause of age-related memory deficits, such that adapting factors like presentation rate or semantic relatedness according to participants' performance generally improves binding memory.

To conclude, we return to the original question regarding improving memory ability in older age. The current results suggest taking advantage of the often-overlooked benefit of getting older: With age comes knowledge, and accordingly, the increased opportunity to relate to-be-learned information to what one already knows.

References

Badham, S. P., Hay, M., Foxon, N., Kaur, K., & Maylor, E. A. (2015). When does prior knowledge disproportionately benefit older adults' memory? *Aging, Neuropsychology, and Cognition*, 1–28. https://doi.org/10.1080/13825585.2015.1099607

Badham, S. P., & Maylor, E. A. (2015). What you know can influence what you are going to know (especially for older adults). *Psychonomic Bulletin & Review*, *22*(1), 141–146. https://doi.org/10.3758/s13423-014-0672-8

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*, 282–293. https://doi.org/10.1037/pag0000317

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(5), P223–P233.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897x00357

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging*, *20*(4), 718–721. https://doi.org/10.1037/0882-7974.20.4.718

Craik, F. I. M., 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

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State - Practical Method for Grading Cognitive State of Patients for Clinician. *Journal of Psychiatric Research*, *12*(3). https://doi.org/10.1016/0022-3956(75)90026-6

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

Heck, D. W., Arnold, N. R., & Arnold, D. (2017). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0869-7

Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attentional resources in the associative deficit of older adults. *Memory & Cognition*, *35*(5), 1162–1174.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & others. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. https://doi.org/10.1037/a0029146

Loaiza, V. M., & Oberauer, K. (2016). Working Memory in Older Age. In N. A. Pachana (Ed.), *Encyclopedia of Geropsychology* (pp. 1–11). https://doi.org/10.1007/978-981-287-080-3_208-2

Loftus, G. R. (1978). On the interpretation of interactions. *Memory & Cognition*, *6*, 312–319.

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–243. https://doi.org/10.1037/a0017619

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., & Balota, D. A. (2009). Aging reduces veridical remembering but increases false remembering: Neuropsychological test correlates of remember–know judgments. *Neuropsychologia*, *47*(11), 2164–2173. https://doi.org/10.1016/j.neuropsychologia.2008.11.025

Mohanty, P. (Pam), Naveh-Benjamin, M., & Ratneshwar, S. (2016). Beneficial effects of semantic memory support on older adults' episodic memory: Differential patterns of support of item and associative information. *Psychology and Aging*, *31*(1), 25–36. https://doi.org/10.1037/pag0000059

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs. (Version 0.9.12-2). Retrieved from http://CRAN.R-project.org/package=BayesFactor

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*, 826–837.

Naveh-Benjamin, Moshe. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1170–1187. https://doi.org/10.1037//0278-7393.26.5.1170

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Oberauer, K. (2005). Binding and Inhibition in Working Memory: Individual and Age Differences in Short-Term Recognition. *Journal of Experimental Psychology: General*, *134*(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

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., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037//0882-7974.17.2.299

R core team. (2017). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428.

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009). *Shipley-2*. Los Angeles, CA: Western Psychological Services.

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: a meta-analysis. *Psychology and Aging*, *10*(4), 527–539.

Umanath, S., & Marsh, E. J. (2014). Understanding How Prior Knowledge Influences Memory in Older Adults. *Perspectives on Psychological Science*, *9*(4), 408–426. https://doi.org/10.1177/1745691614535933

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, *122*(3), 231–249.

Verhaeghen, Paul. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, *18*(2), 332–339. https://doi.org/10.1037/0882-7974.18.2.332

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*(2), 145–160. https://doi.org/10.3758/s13421-011-0158-0

Table 1

Sample characteristics.

	Experiment 1			Experiment 2		
Measure	Younger adults	Older adults	Age Comparison	Younger adults	Older adults	Age Comparison
Age (years)	22.97 (4.69)	70.23 (4.76)	-	20.94 (2.79)	70.62 (4.25)	-
Sex (male/female)	12/18	12/18	$BF_{10} = 0.31$	10/21	13/19	$BF_{10} = 0.37$
MMSE	-	29.00 (0.74)	-	-	28.66 (1.21)	-
Years of education	17.07 (2.63)	15.63 (3.20)	$BF_{10} = 1.16$	16.23 (1.45)	15.19 (3.32)	$BF_{10} = 0.74$
Shipley vocabulary (proportion correct)	0.77 (0.10)	0.93 (0.05)	$BF_{10} = 2.88e+7$	0.72 (0.10)	0.91 (0.05)	$BF_{10} = 2.84e+10$
Number of medications	0.53 (0.90)	2.10 (2.32)	$BF_{10} = 58$	0.52 (0.81)	3.00 (3.14)	$BF_{10} = 601$
Rated current health (1 - 5, 1 = very good)	1.79 (0.86)	1.77 (0.68)	$BF_{10} = 0.08$	2.03 (0.55)	2.06 (0.91)	$BF_{10} = 14.97$
Rated general health (1 - 5, 1 = very good)	1.87 (0.82)	1.73 (0.64)	$BF_{10} = 0.06$	2.03 (0.55)	2.03 (0.93)	$BF_{10} = 3.15$
Rated restrictions of health (1 - 4, 1 = no restrictions)	1.30 (0.53)	1.67 (0.92)	$BF_{10} = 0.12$	1.61 (0.95)	1.81 (0.97)	$BF_{10} = 0.22$

Note. MMSE = mini mental status examination; BF = Bayes factor. Standard deviations in parentheses.

Table 2

Mean performance and 95% highest density intervals (HDIs) sampled from the posterior distributions of the respective age comparison for each observed behavioral measure in Experiments 1 and 2.

		Age group				Age effect			
		Younger adults		Older adults		Effect size			
Exp.	Measure	Mean	95% HDI	Mean	95% HDI	Mean	95% HDI	*p*(ROPE)	BF$_{10}$
1	Proportion of semantically related pairs	0.14	[0.08, 0.21]	0.44	[0.36, 0.51]	1.59	[0.94, 2.25]	0.00	79,973
	WM strict score	0.68	[0.65, 0.71]	0.71	[0.69, 0.73]	0.54	[-0.02, 1.13]	0.05	0.99
	WM lenient score	0.91	[0.89, 0.93]	0.92	[0.90, 0.94]	0.26	[-0.27, 0.78]	0.18	0.40
	EM strict score	0.70	[0.67, 0.74]	0.73	[0.70, 0.75]	0.31	[-0.20, 0.87]	0.15	0.49
	EM lenient score	0.89	[0.88, 0.91]	0.90	[0.88, 0.91]	0.06	[-0.46, 0.60]	0.28	0.27
	EM strict score - corrected	0.78	[0.75, 0.81]	0.75	[0.72, 0.79]	-0.30	[-0.81, 0.23]	0.17	0.46
2	Proportion of semantically related pairs	0.21	[0.12, 0.30]	0.37	[0.29, 0.45]	0.69	[0.14, 1.22]	0.01	4.26
	WM strict score	0.69	[0.67, 0.72]	0.70	[0.68, 0.73]	0.12	[-0.40, 0.63]	0.27	0.28
	WM lenient score	0.90	[0.88, 0.92]	0.90	[0.88, 0.92]	-0.08	[-0.62, 0.46]	0.27	0.26
	EM strict score	0.67	[0.63, 0.71]	0.68	[0.65, 0.70]	0.10	[-0.45, 0.64]	0.26	0.30
	EM lenient score	0.86	[0.84, 0.89]	0.87	[0.85, 0.89]	0.10	[-0.44, 0.58]	0.27	0.28
	EM strict score - corrected	0.69	[0.66, 0.73]	0.66	[0.62, 0.71]	-0.30	[-0.81, 0.22]	0.16	0.50
	EM strict score - untested pairs	0.54	[0.51, 0.58]	0.61	[0.57, 0.65]	0.73	[0.19, 1.27]	0.01	7.51
	EM lenient score - untested pairs	0.79	[0.77, 0.81]	0.82	[0.80, 0.84]	0.53	[-0.01, 1.04]	0.04	1.43

Note. For each effect, the evidence (BF) for the alternative hypothesis over the null is presented (BF$_{10}$). *p*(ROPE) = probability of values within a region of practical equivalence (effect size between -0.1 and 0.1). Exp. = Experiment.

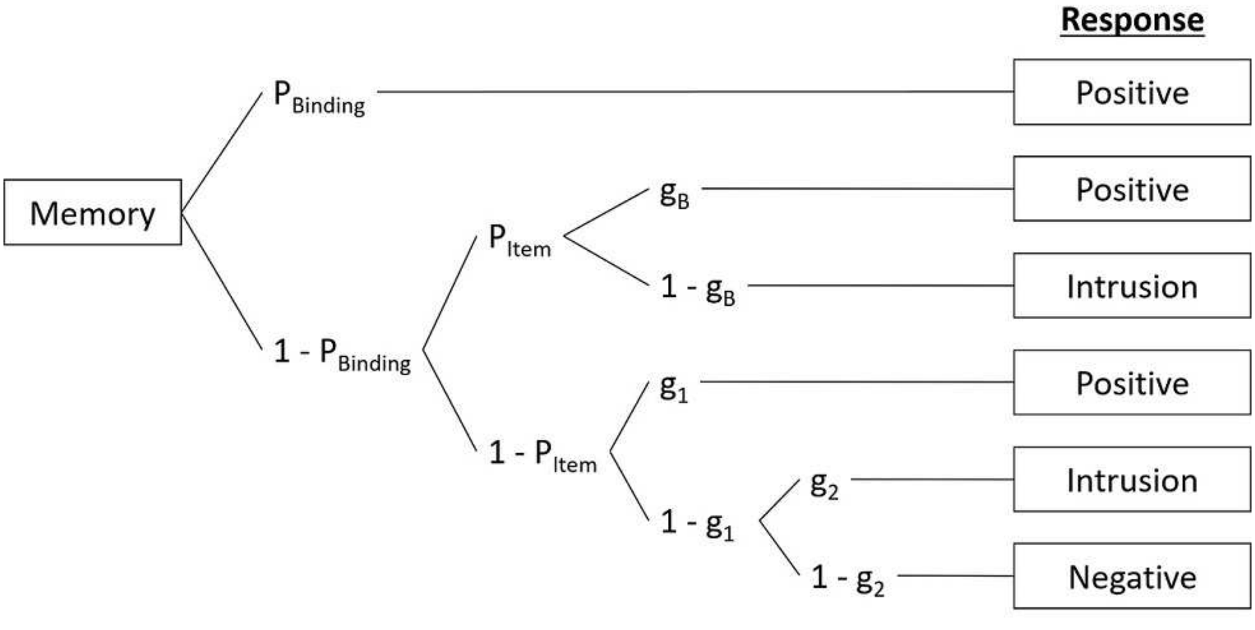

Figure 1. Multinomial processing tree (MPT) model. See text for details.

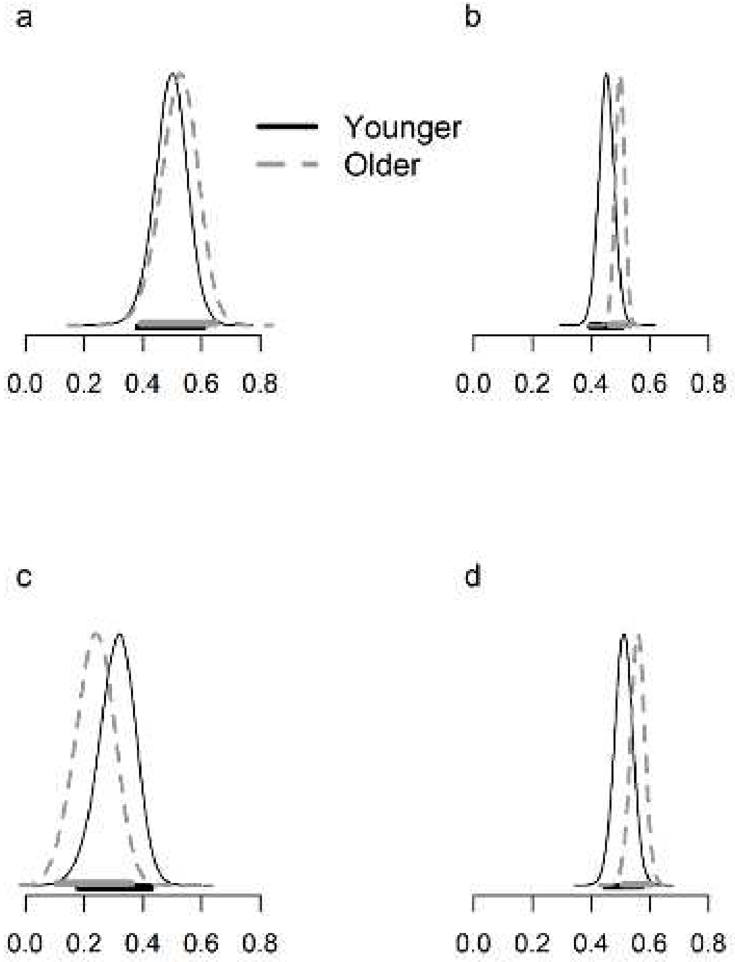

Figure 2. Posterior distributions of the parameters (on probability scale) of the MPTs for younger (solid black curves) and older adults (grey dashed curves) in Experiment 1: Panel a. Item memory in working memory. Panel b. Binding memory in working memory. Panel c. Item memory in episodic memory. Panel d. Binding memory in episodic memory. The bars underneath the curves represent the 95% highest density intervals (HDIs) of the parameters.

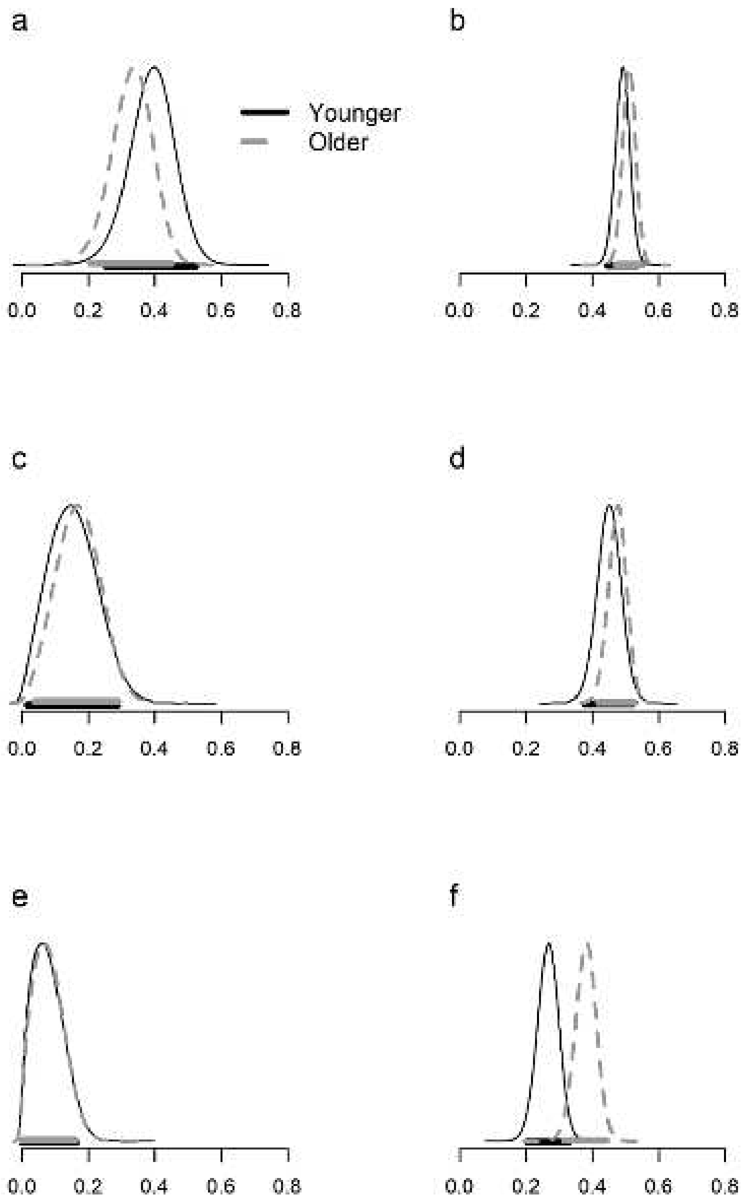

Figure 3. Posterior distributions of the parameters (on probability scale) of the MPTs for younger (solid black curves) and older (grey dashed curves) in Experiment 2: Panel a. Item memory in working memory. Panel b. Binding memory in working memory. Panel c. Item memory in episodic memory – tested pairs. Panel d. Binding memory in episodic memory – tested pairs. Panel e. Item memory in episodic memory – untested pairs. Panel f. Binding memory in episodic memory – untested pairs. The bars underneath the curves represent the 95% highest density intervals (HDIs) of the parameters.