Does limited working memory capacity underlie age differences in associative long-term memory?

Lea M. Bartsch

Vanessa M. Loaiza

Klaus Oberauer

Author Note:

Lea M. Bartsch, Department of Psychology, University of Zurich and University Priority Program Dynamics of Healthy Aging; Vanessa M. Loaiza, Department of Psychology, University of Essex; Klaus Oberauer, Department of Psychology, University of Zurich and University Priority Program Dynamics of Healthy Aging.

We thank Jessica Grub, Gary Hoppeler and Danielle Pessach for helping with data collection. The data and the analysis scripts can be accessed on the Open Science Framework (https://osf.io/n8es5/). The project was funded by the University Priority Program Dynamics of Healthy Aging. Correspondence should be addressed to Lea M. Bartsch, E-mail: l.bartsch@psychologie.uzh.ch. Tel: +41 44 635 74 61

Abstract

Past research has consistently shown that episodic memory (EM) declines with adult age and, according to the associative-deficit hypothesis, the locus of this decline is binding difficulties. We investigated the importance of establishing and maintaining bindings in working memory (WM) for age differences in associative EM. In Experiment 1 we adapted the presentation rate of word pairs for each participant to achieve 67% correct responses during a WM test of bindings in young and older adults. EM for the pairs was tested thereafter in the same way as WM. Equating WM for bindings between young and older adults reduced, but did not fully eliminate, the associative EM deficit in the older adults. In Experiment 2 we varied the set size of word pairs in a WM test, retaining the mean presentation rates for each age group from Experiment 1. If a WM deficit at encoding causes the EM deficit in older adults, both WM and EM performance should decrease with increasing set size. Against this prediction, increasing set size did not affect EM. We conclude that reduced WM capacity does not cause the EM deficit of older adults. Rather, both WM and EM deficits are reflections of a common cause, which can be compensated for by longer encoding time.

Keywords: Episodic long-term Memory, Working Memory, Associative Memory, Bindings, Aging

Introduction

The long-term retention of episodes and events in episodic memory (EM) declines in older age (e.g., Hoyer & Verhaeghen, 2006; Naveh-Benjamin & Old, 2008; Zacks, Hasher, & Li, 2000). So far, the cause of this decline is still under debate. The proposed explanations include reduced speed of processing (Salthouse, 1996), reduced processing resources such us a limited working memory (WM) capacity (Craik & Byrd, 1982), and reduced inhibition (Zacks & Hasher, 1994). The age-related deficit in EM has been characterized as primarily a deficit of old adults in building and retrieving relations (the associative deficit hypothesis, ADH, Naveh-Benjamin, 2000), whereas memory for individual components – referred to as item memory – is relatively intact in older age. The specific age-related decline in associative memory has been shown for various materials including word pairs, picture pairs, and face-name pairs (Bastin & Van Der Linden, 2005; Buchler, et al., 2011; Hara & Naveh-Benjamin, 2015). A meta-analysis evaluating 90 studies on the age-related associative deficit reports large effects sizes of age ($d_A \geq 0.80$) for verbal materials tested with a recognition test (Old & Naveh-Benjamin, 2008). Further, the meta-analysis showed larger age effects on associative than on item memory. Further, it provided evidence that the size of the age-related associative deficit depends, among other variables, on the type of binding formed, with larger deficits for item-item compared to item-context bindings. Furthermore, the age-related associative deficit is larger for recall than recognition test formats (Old & Naveh-Benjamin, 2008). To have a clear characterization of EM decline in older age, it is a priority of cognitive aging researchers to isolate the causes for this disproportionate impairment in associative memory.

Age-related decline may be caused by a working memory deficit

Here we investigated whether WM plays a key role in causing associative deficits of EM in older adults (i.e., the *WM binding deficit hypothesis*). As an alternative, we consider the possibility that age-related deficits in WM and in EM are related through a common cause that impairs WM and EM alike (i.e., the *common cause hypothesis*). We will discuss possible common causes in the General Discussion.

It should be noted that the terminology for relational information varies depending on the memory system. To clarify, for WM relational information is typically referred to as “bindings” whereas for EM the term “associations” is used. We will refer to "bindings" as the general term, encompassing both bindings and associations, but we will continue using the term ‘associative memory deficit’ when referring to the hypothesis introduced under this name.

The WM binding deficit hypothesis starts from the assumption that WM is needed to build and temporarily maintain new bindings, and that WM capacity is a limit on the maintenance of bindings (Oberauer, 2005). According to the WM binding deficit hypothesis, the capacity limit of WM could constrain the bindings formed in EM, and the age-related associative-memory deficit could be a consequence of older adults' reduced WM capacity (e.g., Chalfonte & Johnson, 1996; Hara & Naveh-Benjamin, 2015; Mitchell, Johnson, Raye, Mather, & Esposito, 2000; Park et al., 2002).

To justify the WM binding deficit hypothesis in the first instance, there should be evidence for an age-related deficit for maintaining bindings in WM that is similar to the one shown in EM. The evidence for this assumption is ambiguous. Some studies have provided evidence for an age-related binding deficit in WM (Borg, Leroy, Favre, Laurent, & Thomas-Antérion, 2011; Brown & Brockmole, 2010; Chalfonte & Johnson, 1996; Chen & Naveh-

Benjamin, 2012; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Fandakova, Shing, & Lindenberger, 2013; Mitchell et al., 2000; Oberauer, 2005) whereas others have found no evidence for it (Brockmole, Parra, Della Sala, & Logie, 2008; Brown, Niven, Logie, Rhodes, & Allen, 2017; Parra, Abrahams, Logie, & Della Sala, 2009; Peterson, Schmidt, & Naveh-Benjamin, 2017; Read, Rogers, & Wilson, 2016; Rhodes, Parra, & Logie, 2016).

Assuming that there is an age-related deficit in both maintaining WM bindings and remembering EM bindings, the question remains regarding how they are related: Does the WM binding deficit contribute causally to the associative EM deficit in old age? Support for this notion comes from work varying the study-test retention interval in a continuous recognition task, revealing older adults' binding memory deficit over the short and long term (Chen & Naveh-Benjamin, 2012). However, this result is also consistent with the *common cause hypothesis* according to which binding deficits in WM and more long-term associative-memory deficits are both due to a more general age-related decline in processes that affect memory over the short and the long term. Conclusive evidence for a causal role of impaired WM binding for the age-related associative deficits in EM has not yet been provided. The most convincing evidence for such a role would indicate that experimentally varying any potential age-related binding deficit in WM has strong consequences for the associative deficit in EM. More precisely, if the age-related binding deficit in WM were eliminated, then the corresponding associative deficit in EM should likewise disappear; similarly, if WM bindings were further impaired, then the associative deficit in EM should become larger.

Hara and Naveh-Benjamin (2015) indirectly tested this prediction by simulating older adults’ EM associative deficit in young adults that had to perform a math task with varying difficulty while encoding name–face pairs. Their results showed that young adults’ associative memory declined more than their item memory when engaging in a secondary task during encoding compared to full attention at encoding. This performance pattern mimics the older adults' associative memory deficit. The researchers concluded that the associative deficit in older age is caused by a reduction of their WM resources that may already be exhausted after the individual components are stored but before they are bound together.

One limitation of the study of Hara and Naveh-Benjamin (2015) is that their interpretation relies on an ordinal interaction: Binding memory of the young adults was already worse than their item memory at full attention; under divided attention, this difference increased. This interaction could simply arise because overall performance declines, thereby shifting performance into a more sensitive part of the measurement scale (i.e., further away from ceiling) and consequently amplifying the contrast between item memory and binding memory. This ambiguity in ordinal interactions (Loftus, 1978) prevents any strong interpretation of the study of Hara and Naveh-Benjamin.

Here we take a different approach and test two predictions from the WM binding deficit hypothesis: First, if older adults' WM binding deficit is compensated by giving them more time for encoding the given bindings, this should also compensate the age difference in a subsequent test of EM for the same bindings. This prediction, however, also follows from the assumption that age-related encoding deficits are a common cause of older adults' impaired WM for bindings and their impaired EM for bindings, when the effect of that common cause is compensated for by longer encoding time. The second prediction can adjudicate between these two hypotheses: If increasing the number of items to remember (i.e., the memory set size) impairs the quality of bindings in WM, then increasing set size should likewise impair subsequent EM for the same information in both young and old adults alike. The two experiments of our study tested these two predictions.

Present Study

The goal of the present study was to investigate the importance of establishing and holding bindings in WM to age differences in retention of those bindings in EM. Accordingly, the two reported experiments investigated whether a WM deficit causes the age-related decline in EM. In Experiment 1 we investigated how equating memory for bindings in WM between young and older adults influences older adults' EM for the same bindings. We aimed to equate WM for bindings by adapting the presentation rate of the memoranda according to the subjects' ongoing performance on the WM task, particularly their retention of the bindings, as detailed below. If age-related differences in WM capacity cause the associative-memory deficit in EM, then equating WM binding performance between young and older adults should eliminate the age-related deficit in EM binding. Experiment 2 aimed to test the same hypothesis through a second approach: If WM capacity limits the acquisition of bindings in EM, then increasing the load on WM (i.e., the memory set size) should impair binding memory in a WM test and also in a subsequent EM test for the same bindings.

Measuring Binding and Item Memory

In general, short-term relational recognition tasks require participants to retain bindings between each item (e.g., a word or an object) and another element, such as the item's context (e.g., locations on the screen in which they were presented) or another item (e.g., pairings of words with other words). During the test phase, participants are required to distinguish between the original pairings, recombined pairings, and pairs of new items. Older adults have exhibited more errors on such tasks requiring temporary bindings in WM compared to younger adults, and compared to tasks wherein only an item recognition decision, regardless of the item's bindings to other elements, is required (Fandakova, Shing, & Lindenberger, 2013; Oberauer, 2005). Researchers have subsequently varied these binding tasks to include more types of pairings and stimuli, and modified some details of the test format (De Simoni & von Bastian, 2018; Wilhelm, Hildebrandt, & Oberauer, 2013). For example, Wilhelm and colleagues (2013) presented pairings between two stimuli, such as words and digits, and randomly probed memory for the pairs, with one of the items of the pairing serving as a cue for either its correct match, a completely new item, or an intrusion of an item presented within the trial but not in its correct pairing. This work has provided evidence that a general binding factor represents a common source of variance in typical WM tasks (e.g. complex span, Updating, Recall-1-back; Wilhelm et al., 2013).

Building on the WM binding task of Wilhelm et al. (2013), the WM task in the present study was an immediate memory test in which participants remembered arbitrary word pairs (e.g., *dog – tooth, tree – bottle*) and were tested with a three-alternative forced choice procedure. The test was designed to obtain separate measures of item and binding memory. Specifically, one item from each pair (e.g., *tooth*) was presented with options that included the original correct pairing (e.g., *dog*), a never-presented incorrect item (i.e., a new item; e.g., *book*), or an incorrect lure item that was presented in the trial but not in that pair (e.g., *tree*; see *Figure 1*). EM for the pairs was later tested with the same method used during the WM task. This paradigm allowed separate estimates of binding and item memory for both WM and EM in the same paradigm so that age differences could be investigated without confounding test differences with time of test.

In order to obtain estimates of binding and item memory from the responses in the above task, we applied multinomial process tree (MPT) models to the response frequencies (e.g., Buchner, Erdfelder, & Vaterrodt-Plünnecke, 1995). The structure of the MPT model is shown in *Figure 2*, and is equivalent to a measurement model reported in earlier work (i.e. the independence model, Jacoby, 1999; Jacoby, Debner, & Hay, 2001; see Cooper, Greve, & Henson, 2017 for recent evidence for this approach): The first branch represents whether or not the person correctly remembers the binding of the tested pair. If they remember the binding (with probability *Pb*), they can recollect the correct element previously paired with the cue. If they fail to remember the binding (with probability 1-*Pb*), they can still remember which items have been presented in the current trial (with probability *Pi*), In that case, they can guess between the two items that were in the current trial, leading to a correct response or to a lure response with equal probability (g_b = 0.5). If they fail to recall the items in the trial (with probability 1-*Pi*), they guess among all three response options (correct, lure, new) with equal probability (g_i = 0.333).

The present implementation of the binding task, including the 3-alternative forced-choice set-up together with the MPT measurement model, allows purer estimates of binding and item memory compared to previous paradigms. More precisely, a pure measure of item memory is not achieved by instructing participants to only retain and report on single items – which were nevertheless presented in some context – because incidental encoding of bindings still affects performance (Jaswal & Logie, 2011; Prabhakaran, Narayanan, Zhao, & Gabriel, 2000; Postle, Awh, Serences, Sutterer, & D'Esposito, 2013; Reinitz & Hannigan, 2004; Treisman & Zhang, 2006). This evidence showing that binding memory contaminates many measures of item memory suggests that the aforementioned divergence regarding whether there is a specific age deficit for bindings may at least be partly due to an overestimation of the magnitude of an age difference in item memory. That is, if older adults have a true binding deficit and item memory may be affected by incidental encoding of bindings, then any age difference in item memory may be partly due to the binding deficit even though these are intended to be separate measures. Consequently, this could appear as a more symmetrical age difference between binding and item memory that does not accurately capture the true state of affairs. Thus, our relational recognition binding tasks, together with the MPT measurement model, may better identify the contributions of binding and item memory to performance, which is especially important given the mixed findings of the literature regarding an age-related binding deficit in WM.

In summary, the nature of the relational recognition task and the application of the MPT model allowed us to estimate relatively pure measures of binding and item memory for both WM and EM. Furthermore, the individual and ongoing adaptation of presentation rate of the pairs based on a criterion of correct recollection of bindings in Experiment 1 allowed us to equate WM binding memory between age groups, and to use the resulting presentation rate for Experiment 2. These advantages of the study's design allowed us to distinguish whether equating binding memory in WM between age groups compensates for the age-related associative deficit in EM. Furthermore, if older adults' WM for bindings is impaired more than their WM for items, we expected that equating both age groups with regard to WM for bindings should lead to an age-related *benefit* for item memory (Old & Naveh-Benjamin, 2008). Finally, Experiment 2 utilized the presentation rates approximated in Experiment 1 to assess whether set size similarly impairs binding memory in WM and EM.

Analytic Approach

We used Bayesian statistical analyses, which have been recommended repeatedly for psychological research (e.g., Gallistel, 2009; Rouder, Speckman, Sun, Morey, & Iverson, 2009; Wagenmakers, 2007). Specifically, we implemented hierarchical Bayesian multinomial processing tree (MPT) models. MPT models estimate the probability of latent cognitive states on the basis of categorical data (Batchelder & Riefer, 1999). The hierarchical modeling framework accounts for participant heterogeneity by assuming that the individual parameters are drawn from a distribution describing the population. In this way, the model estimates parameters for each individual, as well as the mean and the dispersion of parameter values in the population. We fit hierarchical MPT models for each age group separately within the *TreeBUGS* Package (Heck, Arnold, & Arnold, 2018) for R (R Core Team, 2017), using the default uniform priors of the package, which are justified in the article by Heck and colleagues (2018). The MPT model of *Figure 2* was applied separately to the responses of each age group in the WM task, and to those of the EM task.

The TreeBUGS package yields Markov-Chain Monte-Carlo (MCMC) samples from the posterior probability distribution of the population mean of the MPT parameters (i.e., estimates of the mean *Pb* and *Pi* for each age group). By subtracting the posterior samples of the young from those of the old adults, we obtained a posterior distribution of the age difference in these parameters. We computed the means and the 95% credibility interval of these differences to assess the effect of age on the MPT parameters (Smith & Batchelder, 2010). The mean of the posterior provides a point-estimate of the effect size (i.e., the central tendency of the posterior difference). The 95% credibility interval gives the smallest range of parameter values over which 95% of the posterior probability is concentrated, and as such provides an assessment of the uncertainty of estimation (i.e., the dispersion of the posterior difference). For inference, we consult the proportion of the posterior probability density of a parameter difference that is larger or smaller than zero; this proportion provides an estimate of the posterior probability that the true effect is positive or negative, respectively.

Experiment 1

Method

Participants

We recruited 30 students (15 female) from the University of Zurich and 30 healthy older adults (15 female) from the Zurich community as participants[1]. They were compensated with either 15 Swiss Francs (about 15 USD) or partial course credit for the one-hour experiment. The study is in line with the rules of the ethics committee of the University of Zurich. Cognitive functioning was screened with the MMSE (Mini-Mental Status Examination; Folstein, Folstein, & McHugh, 1975), indicating age-typical cognitive abilities in our sample of older adults (M = 28.92, SD = 1.07, range = 27 - 30). Table 1 shows the descriptive statistics and posterior distributions of the age effects of our sample. The evidence indicates fewer years of formal education in the older compared with the young adults. The older adults showed better performance than the young adults in a computerized vocabulary test (Mehrfachwahl-Wortschatz Test version B, Lehrl, 2005), consisting of 37 items in which participants are supposed to find an existing word between four similarly sounding non-words. The MWT-B is a marker test for crystallized intelligence. Hence, our sample of young and old adults show typical differences in education and measures of crystallized intelligence (Li et al., 2004).

---

[1] We extended our initial sample of 20 young and 24 older adults during the revision of the manuscript. The use of Bayesian statistics allows for the continuation of sampling (Rouder, 2014; Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, 2017).

Materials and Procedure

The stimuli were randomly drawn from a pool of 589 German concrete nouns for each participant. The nouns were between three and nine letters long and had a mean normalized lemma frequency of 24.76/million (drawn from the dlexdb.de lexical database). Pairs of nouns were created randomly for each participant to serve as the memoranda.

The sequence of an experimental trial is illustrated in *Figure 1*. After the sequential presentation of three word pairs, participants were tested on each pair in random order. Memory was tested with a three-alternative forced-choice task, also illustrated in *Figure 1*: For each probe word, participants selected the word paired with it from three response options: correct, new, and lure (word from another pair). The position of the options on the screen was random, and participants used the mouse to select among them at their own pace. In order to estimate the latent cognitive states of remembering the bindings and items we calculated the number of responses for each of the three response options (correct, lure, and new) for each individual.

The three to-be-remembered word pairs in each trial were sequentially presented from the top to the bottom of the screen. The presentation rate depended on the participants' current cumulative percent of correct binding responses: The adaptive algorithm was a variant of Kaernbach's (1991) weighted up-down algorithm that adjusted the presentation time per pair to achieve performance at 67% *correct* responses (i.e., choice of the correct pairing) for each participant. The algorithm continuously monitored average proportion correct over a moving window of the previous 10 trials. The presentation time for each pair to begin the experiment was set to 1000 ms and 1500 ms for the young and older adults, respectively. For older adults, the presentation rate increased by 200 ms if their moving average performance dropped below 67% correct, and decreased by 100 ms if performance exceeded the criterion. We initially tested 24 young adults with the exact same adaptation method, which unfortunately led to a mean accuracy level higher than we aimed for. We therefore decided to test a new group of young adults, reported here, with stricter adaptation parameters, which theoretically aimed for a 60% criterion, but in practice reached the 67% criterion more closely. More specifically, the presentation rate increased by 180 ms if their moving average dropped below 60% *correct* responses, and decreased by 120 ms if performance exceeded the criterion. For both age groups, the maximum and minimum presentation rates were 5000 ms and 200 ms, respectively.

There were five trials of the WM task in each block. An unrelated distracter task followed each block, in which the participants had to indicate the correctness of presented math equations (e.g. 9 x 8 = 72) for 1 minute. After that followed a delayed cued recognition test in the same format as the immediate test, wherein the participants were probed again with one of the words from each of the 15 pairs from the previous block. The probes were presented in random order, and the probe word as well as the correct response option (i.e., the word actually paired with the probe) were the same as during the WM test. However, the new word among the response options was a completely new word to the experiment (i.e., not the same new word as in the WM test) and the lure word was a word from another pair (i.e., not the same lure word as in the WM test). As during the WM test, the position of the options on the screen was random, and participants used the mouse to select among them at their own pace. This method allowed us to measure binding memory in WM and EM in the exact same format. The experiment comprised five blocks in total.

Results

One older participant's presentation rate had reached the maximum (5000 ms) in the last block; this person was therefore excluded from further analysis, leaving data from 29 older and 30 young adults. For the analysis of the presentation rates as well as the performance during the WM and EM tasks, we also excluded the first block, as we considered it as time for adaptation of the algorithm.

The Bayesian t-test to assess the evidence for the difference of the mean presentation rates per pair revealed decisive evidence for a difference between young (M = 657 ms, SD = 398) and older adults (M = 1724 ms, SD = 932), as the posterior density (PD) of the age effect lies entirely to the left of zero (PD: 0% < 0 < 100%).

The proportion of responses in each of the three response categories (correct, lure, and new items) can be found in Table 2. The critical analysis concerned whether adapting the presentation rate of the word pairs resulted in equated WM for bindings between young and older adults. For this analysis, we compared the age groups with respect to the item-memory and binding-memory parameters derived from the MPT model in *Figure 2*. Figure 3 shows that the adaptation of presentation rates virtually eliminated the age difference in WM binding, as the posterior densities of the WM binding parameter of the young and older adults are overlapping. *Figure 4* depicts the posterior of the age-group difference, showing that the age effect in the mean WM binding parameter is concentrated around zero. We predicted greater item memory in the older adults than the young adults, as their item memory deficit is assumed to be less pronounced than their binding deficit (Mitchell et al., 2000). The difference in the mean WM item parameter supports this hypothesis, as the posterior density lies entirely to the left of zero.

After having ensured equated WM for bindings between age groups, we next examined whether the EM binding deficit was also eliminated. The lower panels of *Figure 3* and *Figure 4* show that the age difference in EM binding parameter still persisted (posterior mean of the age difference = 0.12, highest density interval (HDI) = [0.03, 0.21]). For item memory in EM, the difference in the mean parameter reflects an approximately zero age difference. To ensure that this pattern of results cannot be explained by a mere testing effect (i.e., an advantage of retrieved over non-retrieved information; Rowling, 2014), we also conducted this analysis using EM performance conditionalized on correct WM binding memory. If the pattern is consistent between the former and the conditionalized analyses, then the age deficit in EM bindings is unlikely to be attributable to any differential retrieval practice that the pairs received in WM. The evidence for a remaining age difference in the EM binding parameter persisted in this analysis (see *Figure 5*). In order to quantify the remaining EM binding deficit, we calculated the effect sizes the same way as in the meta-analysis (Old & Naveh-Benjamin, 2008), from our group-mean posterior parameter estimates and the respective sigma's. The resulting distribution of effect sizes had a mean of d = 0.041 (HDI = [0.01, 0.08]), with 99.6% of the posterior density above zero.

Discussion

The goal of Experiment 1 was to investigate the importance of establishing and holding bindings in WM to age differences in EM bindings. We successfully equated WM for bindings between age groups by adapting the presentation rate of to-be-remembered word pairs. Our results show that this did not eliminate the EM deficit in old age. This implies that older adults' EM deficit is not entirely caused by a WM deficit at encoding. Nevertheless, by compensating the WM deficit, we substantially reduced the age-related EM deficit in comparison to previous studies (d = 0.04 compared to the meta-analysis of age differences in EM with effect sizes of d >.80 for verbal material and recognition tests, Old & Naveh-Benjamin, 2008). Therefore, the EM deficit could still be in part due to a binding deficit in WM, as the remaining binding deficit in EM was rather small. Alternatively, the results of Experiment 1 could be explained by the *common cause hypothesis:* Age-related deficits in WM and EM could be reflections of a common cause, which is partly compensated for by longer encoding time, leading to the reduction in EM binding deficits. For instance, both forms of memory might suffer from a similar age-related slowing of consolidation, the hypothetical process converting fragile, transient representations into more stable memory representations (Chun & Potter, 1995; Jolicœur & Dell'Acqua, 1998; Wixted, 2004).

The goals of Experiment 2 were to test whether the EM associative deficit is in part caused by a binding deficit in WM, and to distinguish that hypothesis from the *common cause hypothesis*. According to the *WM deficit hypothesis*, interference between the word pairs in WM causes them to be represented less precisely or less robustly in older adults than in young adults, thereby leading to impaired EM representations. Accordingly, increasing the number of word pairs (i.e., set size) in WM should have a corresponding detrimental effect on EM. We would therefore predict that increasing memory set size leads to poorer performance not only in an immediate WM test but also in a subsequent test of EM. Moreover, because WM capacity declines in older age (Hale et al., 2011), we would predict an interaction between set size and age group on EM, such that older adults should show worse EM performance than young adults, especially as set size increases from a low value (at which both age groups' WM capacity is still sufficient to maintain all bindings well) to a value at which the WM of an average young person can still accommodate all bindings whereas the WM of older adults begins to struggle. The *common cause account* of the findings of Experiment 1, by contrast, predicts that the critical factor for EM performance is the time given to attend to and process the individual pairs, independent of how many other pairs are held in WM concurrently. In this case, increasing set size while keeping the presentation rate per pair constant should have no effect on EM binding. In order to adjudicate between these accounts, in Experiment 2 we varied set size of the tested pairs in WM while holding constant the presentation rate of the pairs at that for which WM for bindings of three pairs was equal between young and older adults in Experiment 1.

Experiment 2

Method

Participants

We recruited a new sample of 30 students (15 female) from the University of Zurich and 30 healthy older (15 female) adults from the community as participants. They were compensated with either 15 Swiss Francs (about 15 USD) or partial course credit for the one-hour experiment. The study is in line with the rules of the ethics committee of the University of Zurich. Cognitive functioning was screened with the MMSE (Folstein et al., 1975), indicating normal cognitive abilities in the sample of older adults ($M = 28.82$, $SD = 1.47$, range = 27 – 30). As evident by the posterior densities of the age differences in Table 1, the older adults had completed fewer years of formal education than the young adult and showed better performance in a computerized version of the MWT-B vocabulary test (Lehrl, 2005) than the young adults.

Materials and Procedure

The materials and procedure for Experiment 2 were very similar to Experiment 1. The principal differences were the following: Set size was varied across trials by presenting between

2 and 6 word pairs per trial. As in Experiment 1, the word pairs were sequentially presented from the top to the bottom of the screen. Furthermore, presentation rate was held constant within each age group at the mean presentation rates of the first 20 young and 24 older adults from Experiment 1, at which young and older adults showed equated WM for bindings for a set of three pairs (young = 710 ms and older adults = 1760 ms)[2]. As in Experiment 1, memory for each pair was probed immediately and after a delay. There was one trial of each set size per block, and seven blocks in the experiment.

Results

We analyzed the number of correct, lure, and new item responses with the same hierarchical MPT model as in Experiment 1 using the *TreeBUGS* package in the R environment. We applied separate MPT models for WM and EM, as well as for each set size level and age group. The proportions of responses in the three response categories (correct, lure, and new items) are shown in *Figure 6*. The critical analysis concerned whether increasing interference in WM through increased set size decreases EM for bindings, and if so, whether that decrease was more pronounced in older than young adults even when using the presentation rates for which young and older adults showed equivalent WM binding at one of the lower set sizes (3 pairs) in Experiment 1. Figure 7 shows the posterior estimates for the main effect of set size, as well as the interaction effect of set size by age, for the parameters of the MPT models.

Unsurprisingly, increasing the number of to-be-remembered pairs in a trial reduced the WM binding performance for those pairs, represented by the change in parameter *Pb* (PD = 0% <

[2] These presentation times were derived from the average presentation time per age group from the initial sample of 20 younger and 24 older adults. The mean presentation times reported for

0 < 100%). Furthermore, as indicated by the interaction effect, the age-related difference varied with set size: young adults showed worse WM binding performance than older adults at set size 4, 5 and 6 (PD: 98.7% < 0 < 1.3%, PD = 89.6% < 0 < 10.4%, PD: 99% < 0 < 1%, respectively), but better WM binding performance than older adults at set size 2 (PD: 5.3% < 0 < 94.7%). For set size 3, the posterior for the age difference was centered on zero (PD: 41.3% < 0 < 58.7%), replicating the finding from Experiment 1 of approximately equivalent WM binding performance in both age groups at this set size with the given presentation rates.

The analysis of the parameters for item memory in WM revealed no effect of set size, neither for young (PD = 29.2% < 0 < 70.8%) nor for older adults (PD = 36.3% < 0 < 63.7%). For the main effect of age, 99.7% of the posterior density lay to the left side of zero, implying – as in Experiment 1 – an age-related *benefit* for item memory in WM. This means that, after compensating for older adults' difficulty with maintaining bindings in WM, their item memory was better than that of young adults.

The analysis of the binding parameter for EM revealed evidence that the age difference was rather small, as the posterior density included considerable proportions on both sides of zero (PD: 16.5% < 0 < 83.5%). The critical analysis concerned whether set size affected EM in a similar way as it affected WM performance of young and old adults. There was, if anything, a very small main effect of set size, as the posterior density included considerable proportions on both sides of zero (PD: 86.2% < 0 < 13.8%). The difference in parameters between the age groups at each set size are shown in Figure 7C. There was evidence for an interaction between set size and age (PD: 2.1% < 0 < 97.9%), driven by better EM for bindings of young adults at set size 2 & 3 ($PD_{setsize2}$: 6.5% < 0 < 93.5%, $PD_{setsize3}$: 7.8% < 0 < 92.2%), whereas EM for bindings

---

Experiment 1, include additional subjects and are therefore slightly different.

was equivalent between age groups at larger set sizes (PD $_{setsize4}$: 47.8% < 0 < 52.2%, PD $_{setsize5}$ = 36% < 0 < 64%, PD $_{setsize6}$: 64.8% < 0 < 35.2%). As for Experiment 1 we calculated the distribution of effect sizes from the estimates of the group-mean differences for the binding parameter at each set size. The analysis revealed small effect sizes for the age differences of the binding parameter at set sizes 2 and 3 (mean PD of $d_{setsize2}$= 0.02, HDR = [-0.01, 0.05]; mean PD of $d_{setsize3}$ = 0.02 HDR = [-0.01, 0.05]. For set sizes 4, 5, and 6 the posterior densities of the effect sizes were distributed around zero (mean PD of $d_{setsize3}$= 0, HDR = [-0.03, 0.03]; mean PD of $d_{setsize4}$= 0.01, HDR = [-0.02, 0.03]; mean PD of $d_{setsize5}$= -0.01, HDR = [-0.03, 0.02]).

For the parameter of item memory in EM, the analysis revealed no evidence for a main effect of set size (PD: 62.6% < 0 < 37.4%), nor a main effect of age (PD: 26.8% < 0 < 73.2%), and no evidence for an interaction between them (PD: 64% < 0 < 36%).

As in Experiment 1, we further ensured that the pattern of results of Experiment 2 was not attributable to a mere testing effect. To this end, we analyzed the EM performance also conditionalized on whether or not the pairs were correctly remembered during the WM test. The analysis confirmed the negligible age deficit in EM binding and item memory when performance was conditionalized on accurate WM binding (binding memory: PD: 62.6% < 0 < 37.4%, item memory: PD: 58.7% < 0 < 41.3%). Furthermore, the conditionalized analysis similarly showed negligible set-size effects for item memory (PD: 64% < 0 < 36%) and evidence for a set size effect for binding memory (PD: 2.1% < 0 < 97.9%).

Discussion

To summarize, as a successful manipulation check, in Experiment 2 we replicated the equated binding performance in WM between young and older adults at set size three with the presentation rates from Experiment 1. Also, we replicated the substantive finding of Experiment 1, namely, that the retention of the bindings in EM was better in young than in older adults at set size three, despite the age-related compensation in WM. Furthermore, set size had the expected detrimental effect on WM bindings in young and older adults.

Despite its detrimental effect on both age groups' WM for bindings, increases in set size had no such effect on EM for either age group. Instead, the findings showed that the small age-related differences of EM bindings at set sizes 2 and 3 disappeared at the larger set sizes, as older adults' EM for bindings slightly increased at higher set sizes, leaving no evidence for an impact of age on bindings in EM.

From the *WM deficit hypothesis,* we predicted a set-size effect not only on WM but also on EM. In addition, we predicted an interaction between set size and age group in EM, such that older adults should show worse performance than young adults particularly at higher set sizes The above findings refute both predictions, decisively ruling out the *WM deficit hypothesis.* Our finding that old adults' WM deficit can be compensated for by giving them longer time for encoding and consolidating the memory pairs, and that this largely (Experiment 1) or entirely (Experiment 2) compensated for their EM deficit, is better explained by the *common cause hypothesis.* Specifically, older adults might be slower in consolidating information in both WM and EM, and this slowing is partially responsible for their reduced binding ability in tests of WM as well as EM.

Experiment 2 yielded one unexpected effect: The set-size effect on WM bindings was larger for young adults, resulting in worse WM performance relative to older adults at larger set sizes, and somewhat better performance at the smallest set size. This result is surprising in light of a recent study by Read (2016) showing that increases in set size similarly impaired feature-location and feature-feature bindings in younger and older adults. That said, our results are in line with findings from Boujut & Clarys (2016). We can only offer a speculative post-hoc explanation for this interaction of set size with age: We compensated the age-related WM binding deficit by giving older adults substantially more time for encoding each pair. We tentatively concluded that this time is used to consolidate bindings better in both WM and EM. Perhaps the longer presentation time is used primarily for establishing better memory representations in EM, with relatively little effect on WM. In addition to improving delayed memory, better EM representations could also assist performance in the immediate test (intended to measure WM). As a consequence, older adults' performance on the WM test would rely more strongly on EM than that of younger adults. As EM is not affected by set size, this would result in a flatter set-size effect in the WM-test performance of older compared to younger adults.

General Discussion

The goal of the present study was to investigate the importance of maintaining bindings in WM for age-related EM deficits, especially the disproportionate associative deficit in older age. Using a novel paradigm that adapted the presentation rate of word pairs for young and older adults, we equated WM for bindings and subsequently observed a small but persistent EM binding deficit in older adults in Experiment 1. Further, the results of Experiment 2 were incompatible with the *WM deficit hypothesis*, suggesting instead that inefficiency at encoding or during consolidation of memory traces may cause a more general age-related deficit in retaining bindings in WM and EM alike, in line with the *common cause hypothesis*.

The finding that memory set size, although strongly affecting WM performance, had no effect on subsequent EM for the same information for either age group contradicts the *WM deficit hypothesis* that has been advanced in previous work. For example, Hara & Naveh-

Benjamin (2015) simulated the age-related associative deficit in EM by having young adults encode materials under divided attention, and they interpreted their result as consistent with the *WM deficit hypothesis*, such that an associative deficit in WM causes EM associative deficits in older adults. The present investigation questions this claim and suggests an alternative explanation: Instead of simulating a WM deficit, the divided attention manipulation reduced the time available for encoding, and therefore impaired young adults' EM to a similar extent as a naturally occurring encoding deficit of old adults.

Although the current results rule out the hypothesis that WM binding deficits cause EM deficits in older adults, they leave us with a new question: Which process did older adults in our experiments engage during their longer encoding time to reduce their deficit in both the WM and the EM tests? One possibility could be that older adults invested the increased encoding time to use (more) normatively effective strategies, such as elaboration. Findings from Bailey, Dunlosky, and Hertzog (2009) speak against this notion: These researchers showed that young and older adults report a similar prevalence of normatively effective strategies during WM tasks. In contrast, measures of processing speed accounted for a substantial proportion of the age-related variance in WM performance.

Our findings are consistent with the general slowing hypothesis (Salthouse, 1996; for an overview see Hartley, 2006), which proposes reduced processing speed to account for age-related differences in cognitive functions. The general slowing hypothesis emerged from consistent observations that older participants show longer reaction times to respond to stimuli, which supposedly represents slowing of perceptual, motor, and cognitive processes. Although slowing as common cause for age-related deficits in many tasks is attractive for its parsimony, it has long been debated what actually causes the phenomenon. For example, one could interpret the present results as consistent with the slowing of consolidation, such that older adults differentially struggle to create stable memory representations and require more time to do so compared to younger adults, thereby causing binding deficits in WM and EM. However, the current study cannot dissociate whether general cognitive slowing or a more specific slowing of processes such as consolidation cause the age-related binding deficit.

Furthermore, although our results are in line with the general slowing hypothesis, they do not strongly support this interpretation because other interpretations are just as plausible. For instance, it has been shown that older adults suffer from more neural noise, and therefore create less distinctive representations between successively presented pairs (Noack, Lövdén, & Lindenberger, 2012). As a result, the pairs are encoded with more overlapping representations, and at retrieval, the probe cues other words in the trial in addition to the actual target. This would cause binding memory impairments while leaving intact memory for items. For the current study, longer processing time could have led to more distinct, less noisy representations given the greater temporal separation of the pairs. This would result in better distinctiveness of the material, thereby reducing the binding deficit of the older adults. At larger set sizes the additional time may be particularly useful to engage in differential encoding (i.e., forming representations of the word pairs in the memory set that emphasize the differences between them). Accordingly, what may appear at first glance as a general slowing deficit could instead reflect more time to engage in specific processes that may be deficient in older age, such as greater use of normatively effective strategies, consolidation of traces into stable representations, and reduction of representational overlap.

The current research is also relevant to previous work that has considered variation of presentation time to examine age deficits in WM. For example, Oberauer and Kliegl (2001) applied an adaptive algorithm (Kaernbach, 1991) similar to our Experiment 1 to vary presentation rates for young and older subjects in a WM updating task. They showed that WM capacity limits of the old adults could not be fully compensated by increasing encoding and updating times, as the young adults benefited from longer times too, and reached a higher asymptotic performance level with increasing time for each updating step. These findings indicate that age deficits in binding cannot be solely attributed to slower encoding or consolidation. Other work has similarly tried to compensate for the age-related memory deficit by increasing encoding time but has not managed to fully do so. A study by Sander and colleagues (Sander, Werkle-Bergner, & Lindenberger, 2011) showed WM performance of older adults increased with longer presentation rates; nevertheless, the older adults did not reach the level of the young adults' performance. The failure to fully compensate the age-related WM deficit could have occurred because they chose fixed longer presentation rates for older compared to younger adults, rather than adaptively varying presentation rates as in the current study. The choice of the presentation rate for older adults might just not have been slow enough to fully compensate their WM deficit. Two further recent studies investigated the effects of encoding time on WM bindings in young and old adults: Rhodes et al. (2016) found no differential effect of longer (2500ms) compared to shorter (900ms) encoding time on WM binding performance of older adults. Similarly, Brown et al. (2017) found that although older adults profited more than young adults from a longer (1500ms vs. 900ms) encoding time overall, this did not differentially effect binding over item memory. The findings of these studies do not conflict with our results as these studies' procedures allowed both age groups a fixed amount of additional time at encoding, which also permits the young adults to improve their bindings in WM. This is different to our approach here, where we calibrated the encoding time to *compensate* the older adults' lower performance. Another difference between the present study and that of Brown et al. (2017) could be the use of different stimuli (i.e., binding of shape and color vs. pairs of unrelated words): meta-analytic evidence suggests the age-related associative deficit is smaller for verbal compared with visuospatial materials (Old & Naveh-Benjamin, 2008). Accordingly, visuospatial memoranda may require an even larger adjustment of encoding time for older adults to compensate for their relatively larger binding deficit.

In summary, the present study tested the causal role of WM for the age-related binding deficit in EM. The results ruled out the *WM deficit hypothesis* that asserts that the binding deficit is due to a deficit to establish and maintain bindings in WM. Instead, the evidence was congruent with a common cause of both deficits. One plausible candidate for this common cause lies in less efficient encoding and consolidation processes in older age.

References

Bailey, H., Dunlosky, J., & Hertzog, C. (2009). Does Differential Strategy Use Account for Age-Related Deficits in Working-Memory Performance? *Psychol Aging*, *24*(1), 82–92. https://doi.org/10.1037/a0014078

Bastin, C., & Van Der Linden, M. (2005). The effects of aging on the recognition of different types of associations. *Experimental Aging Research*, *32*(1), 61–77. https://doi.org/10.1080/03610730500326291

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57–86.

Borg, C., Leroy, N., Favre, E., Laurent, B., & Thomas-Antérion, C. (2011). How emotional pictures influence visuospatial binding in short-term memory in ageing and Alzheimer's disease? *Brain and Cognition*, *76*(1), 20–25. https://doi.org/10.1016/j.bandc.2011.03.008

Boujut, A., & Clarys, D. (2016). The effect of ageing on recollection: the role of the binding updating process. *Memory*, *24*(9), 1231–1242. https://doi.org/10.1080/09658211.2015.1091893

Brockmole, J. R., Parra, M. A., Sala, S. D., & Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin & Review*, *15*(3), 543–547. https://doi.org/10.3758/PBR.15.3.543

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology*, *63*(10), 2067–2079. https://doi.org/10.1080/17470211003721675

Brown, L. A., Niven, E. H., Logie, R. H., Rhodes, S., & Allen, R. J. (2017). Visual feature binding in younger and older adults: encoding and suffix interference effects. *Memory*, *25*(2), 261–275. https://doi.org/10.1080/09658211.2016.1156705

Buchler, N. G., Faunce, P., Light, L. L., Gottfredson, N., & Reder, L. M. (2011). Effects of repetition on associative recognition in young and older adults: Item and associative strengthening. *Psychology and Aging*, *26*(1), 111–126. https://doi.org/10.1037/a0020816

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General*, *124*(2), 137.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416. https://doi.org/10.3758/BF03200930

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*(3), 666–682. https://doi.org/10.1037/a0026943

Chun, M. M., & Potter, M. C. (1995). A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology*, *21*(1), 109–127.

Cooper, E., Greve, A., & Henson, R. N. (2017). Assumptions behind scoring source versus item memory: Effects of age, hippocampal lesions and mild memory problems. *Cortex*, *91*, 297–315. https://doi.org/10.1016/j.cortex.2017.01.001

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, *42*(6), 1089–1102. https://doi.org/10.1037/0012-1649.42.6.1089

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits. In *Aging and cognitive processes* (pp. 191–211). Springer.

De Simoni, C., & von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer, (April). https://doi.org/10.1037/xge0000453

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013). Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. *Developmental Psychology*, *49*(10), 1822–32. https://doi.org/10.1037/a0031361

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*(2), 439–453. https://doi.org/10.1037/a0015251

Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging*, *26*(1), 92–110. https://doi.org/10.1037/a0021483.The

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

Hartley, A. (2006). Changing role of the speed of processing construct in the cognitive psychology of human aging. *Handbook of the Psychology of Aging*, 183–207.

Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, *50*(1), 264–284. https://doi.org/10.3758/s13428-017-0869-7

Hoyer, W. J., & Verhaeghen, P. (2006). *Memory aging*. *Handbook of the psychology of aging* (Vol. 6).

Jacoby, L. L. (1999). Ironic Effects of Repetition: Measuring Age-Related Differences in Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(1), 3–22. https://doi.org/10.1037/0278-7393.25.1.3

Jacoby, L. L., Debner, J. A., & Hay, J. F. (2001). Proactive Interference, Accessibility Bias, and Process Dissociations: Valid Subjective Reports of Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*(3), 686–700. https://doi.org/10.1037/0278-7393.27.3.686

Jaswal, S., & Logie, R. H. (2011). Configural encoding in visual feature binding. *Journal of Cognitive Psychology*, *23*(5), 586–603. https://doi.org/10.1080/20445911.2011.570256

Jolicœur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Lehrl, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B* (5th ed.). Balingen: Spitta Verlag.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual cognitive processes across the life span. *Psychological Science*, *15*(3), 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*(3), 312–319. https://doi.org/10.3758/BF03197461

Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., & Esposito, M. D. (2000). Aging and Reflective Processes of Working Memory: Binding and Test Load Deficits. *Psychology and*

Aging, *15*(3), 527–541. https://doi.org/10.1037/70882-

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. *R Package Version 0.9*, *9*.

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*(5), 1170–1187. https://doi.org/10.1037/0278-7393.26.5.1170

Naveh-Benjamin, M., & Old, S. R. (2008). 2.41 Aging and Memory. *Learning and Memory: A Comprehensive Reference*, 787–808. https://doi.org/http://dx.doi.org/10.1016/B978-012370509-9.00133-9

Noack, H., Lövdén, M., & Lindenberger, U. (2012). Normal aging increases discriminal dispersion in visuospatial short-term memory. *Psychology and Aging*, *27*(3), 627–637. https://doi.org/10.1037/a0027251

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

Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology*, *13*(1-2), 187–215. https://doi.org/10.1080/09541440042000278

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

Parra, M. A., Abrahams, S., Logie, R. H., & Sala, S. Della. (2009). Age and binding within-dimension features in visual short-term memory. *Neuroscience Letters*, *449*(1), 1–5. https://doi.org/10.1016/j.neulet.2008.10.069

Peterson, D. J., Schmidt, N. E., & Naveh-Benjamin, M. (2017). The role of schematic support in age-related associative deficits in short-term and long-term memory. *Journal of Memory and Language*, *92*, 79–97. https://doi.org/10.1016/j.jml.2016.05.007

Postle, B. R., Awh, E., Serences, J. T., Sutterer, D. W., & D'Esposito, M. (2013). The positional-specificity effect reveals a passive-trace contribution to visual short-term memory. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0083483

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria; 2014. Retrieved from https://www.r-project.org/.

Read, C. A., Rogers, J. M., & Wilson, P. H. (2016). Working memory binding of visual object features in older adults. *Aging, Neuropsychology, and Cognition*, *23*(3), 263–281. https://doi.org/10.1080/13825585.2015.1083937

Rhodes, S., Parra, M. A., & Logie, R. H. (2016). Ageing and feature binding in visual working memory: The role of presentation time. *Quarterly Journal of Experimental Psychology*, *69*(4), 654–668. https://doi.org/10.1080/17470218.2015.1038571

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301–308.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Diff erences in Cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Sander, M. C., Werkle-Bergner, M., & Lindenberger, U. (2011). Binding and strategic selection in working memory: A lifespan dissociation. *Psychology and Aging*, *26*(3), 612–624. https://doi.org/10.1037/a0023055

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322.

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology*, *54*(1), 167–183.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society*, *64*(4), 583–639.

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719. https://doi.org/10.3758/BF03195932

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, *4*(JUL), 1–22. https://doi.org/10.3389/fpsyg.2013.00433

Wixted, J. T. (2004). The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology*, *55*(1), 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and* *language* (pp. pp. 241–264). San Diego, CA: CA: Academic Press.

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human Memory. *The Handbook of Aging and Cognition*, 293–357. https://doi.org/10.3758/s13423-011-0086-9

Table 1 Sample Description (means (and standard deviations)) of Experiment 1 and 2

Experiment	Age Group	Age	years of education	vocabulary
1	Younger	24.06 (3.77)	14.70 (3.07)	77.07 (13.22)
	Older	71.26 (3.99)	13.81 (3.40)	86.29 (3.43)
	$PD_{age\text{-}effect}$	-	16.9% < 0 < 83.1%	99.9% < 0 < 0.1%
2	Younger	24.84 (2.94)	15.84 (3.08)	77.90 (14.26)
	Older	71.07 (4.08)	13.66 (4.54)	85.27 (6.95)
	$PD_{age\text{-}effect}$	-	3.1% < 0 < 96.9%	98.7 % < 0 < 1.3%

Note. The posterior density (PD) of the age effects. Zero represents the point of no age differences, and the percentages indicate how much of the estimated effect's posterior distribution lies below and above 0. Values below 0 reflect an advantage of older adults whereas positive values indicate a younger adults advantage.

Table 2 Percent of responses per category in Experiment 1.

Memory test	Age Group	Correct	Lure	New
WM	Younger	66.47 (4.36)	20.97 (4.11)	12.56 (2.88)
	Older	69.14 (7.04)	23.71 (7.4)	7.14 (4.41)
EM	Younger	61.81 (9.7)	25.61 (7.18)	12.58 (4.33)
	Older	55.43 (8.94)	31.71 (8.57)	12.86 (7.13)

Note. The standard deviation is marked in parentheses.

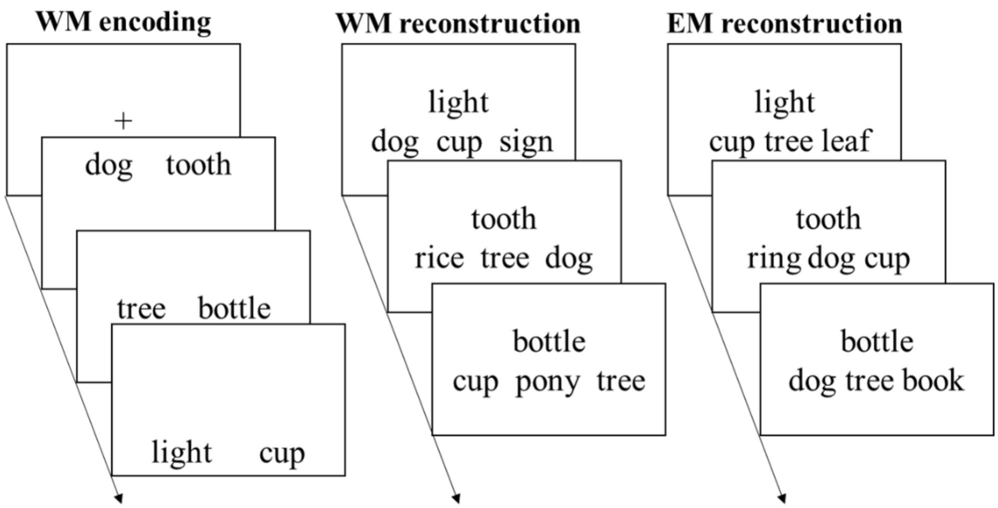

Figure 1 A representation of the paradigm used in Experiments 1 and 2.

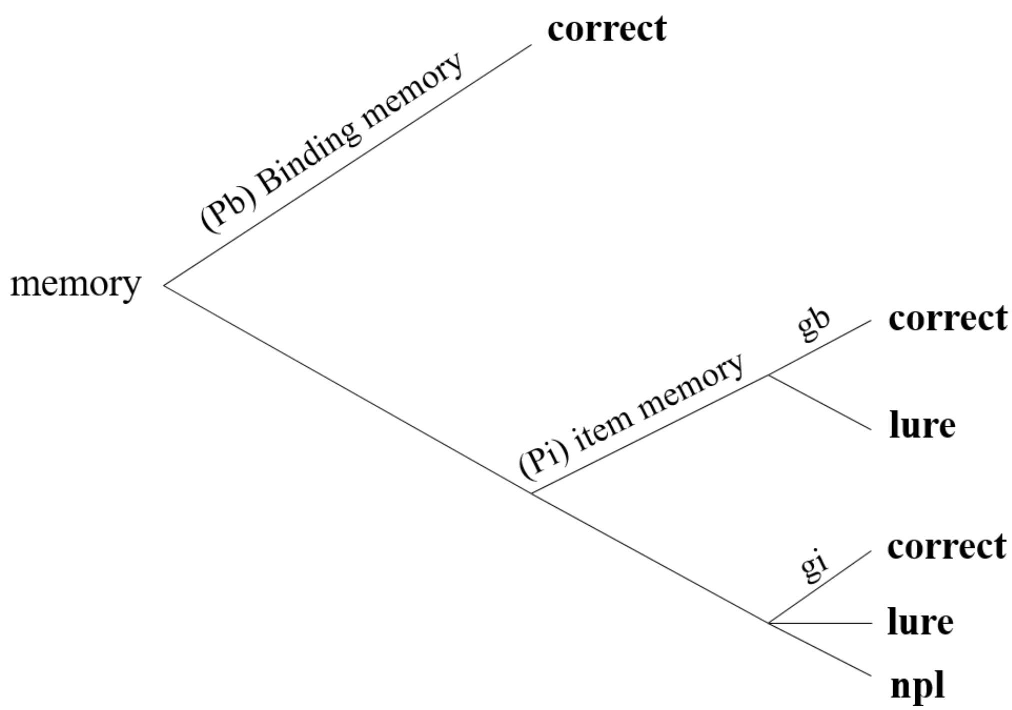

Figure 2 Multinomial-process tree (MPT) model for memory of bindings in Experiments 1 and 2. See section *Measuring Binding and Item Memory* for details.

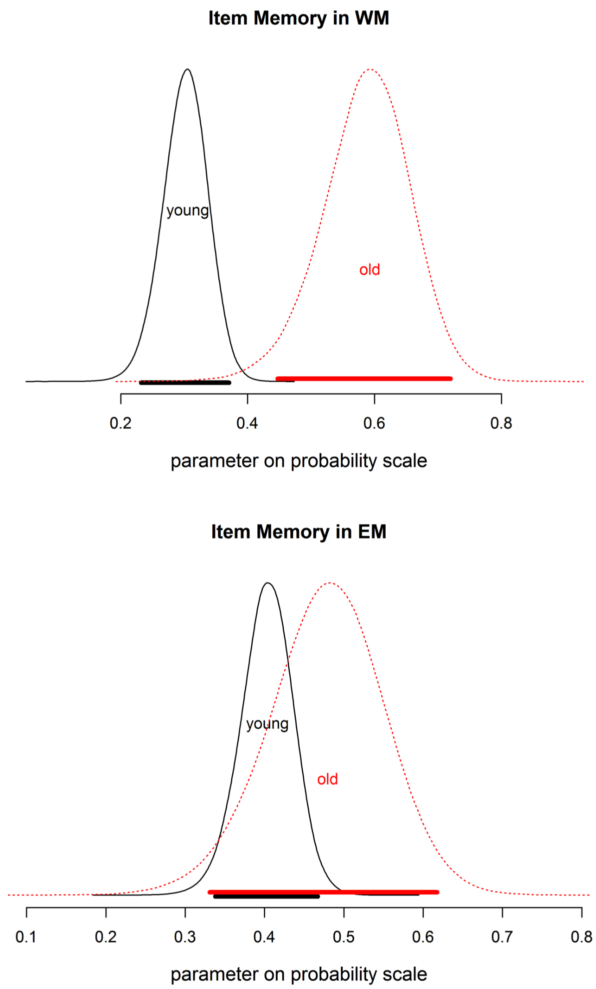

Figure 3 Posterior distributions of the parameters of the MPTs for young and older adults of Experiment 1. The horizontal lines represent the 95% highest density intervals.

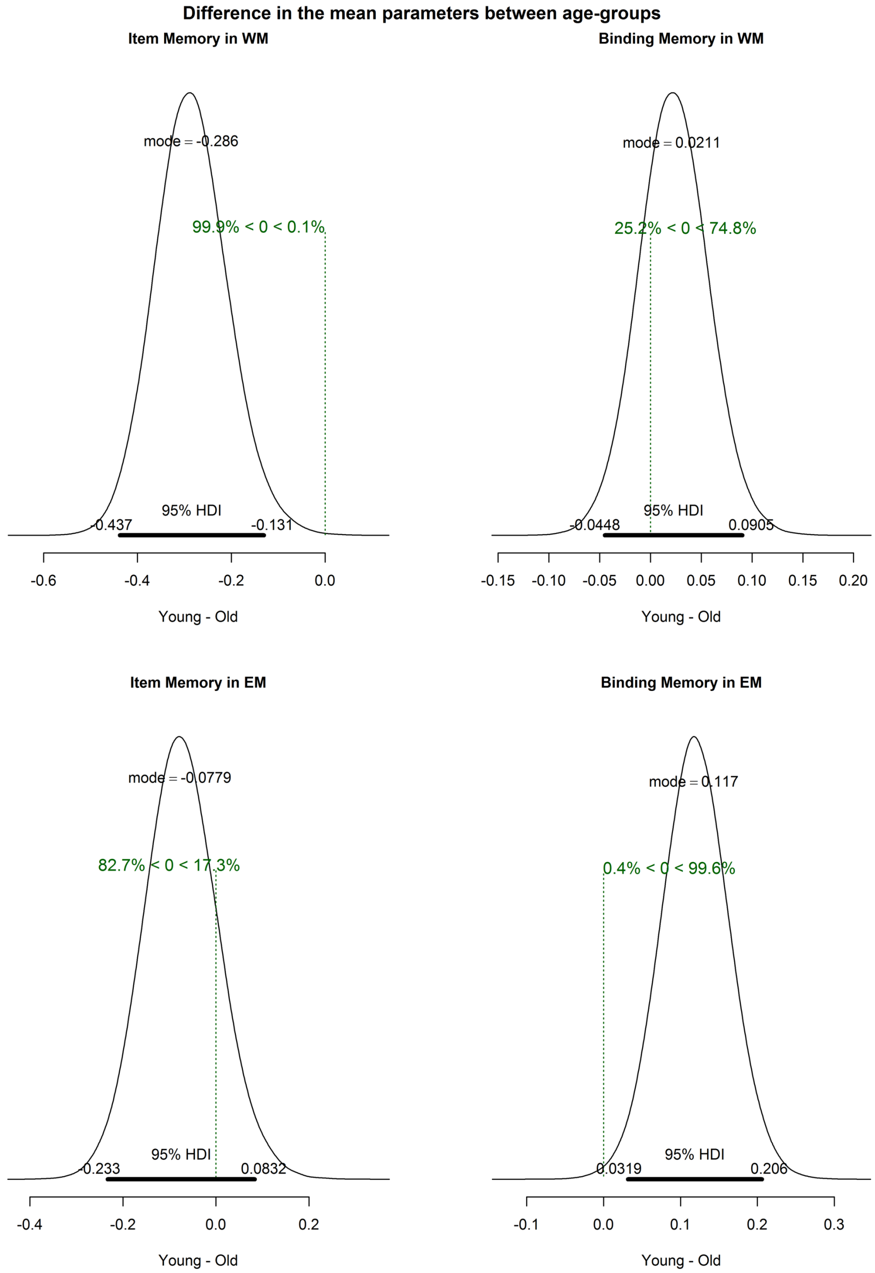

Figure 4 Posterior distributions of differences of mean the parameters between the age groups of Experiment 1. The mode with its respective highest density intervals reflect the effect size of any age difference. The dotted line indicates the point of no age differences, and the percentages indicates the credibility interval of the difference. Values below 0 reflect an advantage for the older adults.

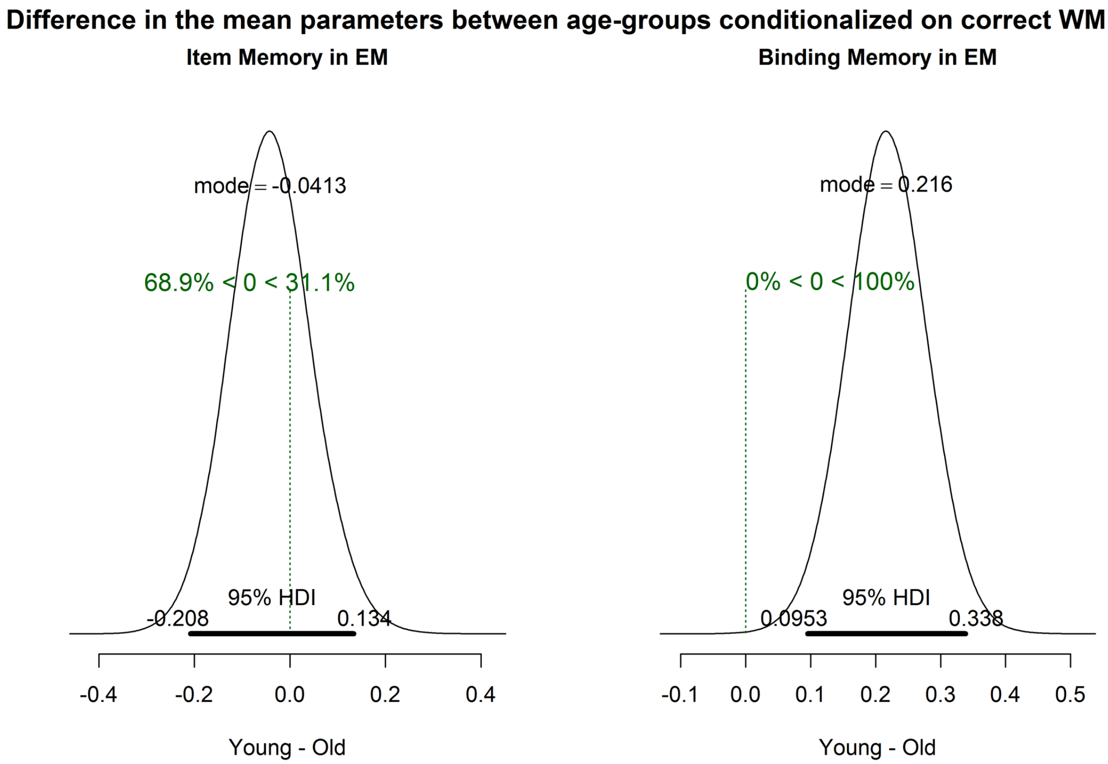

Figure 5 Posterior distributions of differences of mean the parameters between the age groups conditionalized on correct binding memory in WM of Experiment 1. The dotted line indicates the point of no age differences, and the percentages indicates the credibility interval of the difference. Values below 0 reflect an advantage for the older adults.

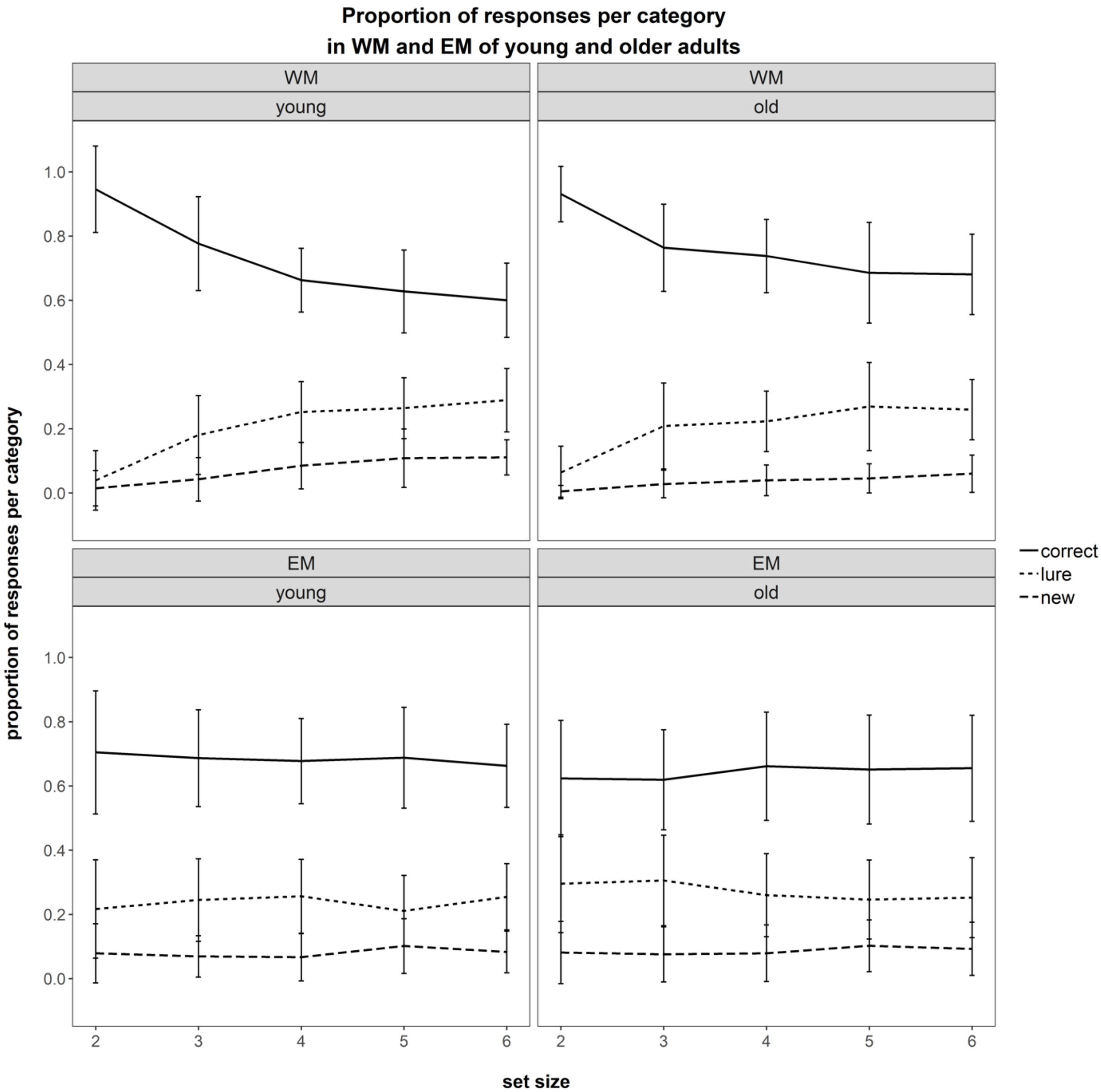

Figure 6 Proportion of responses per category in WM and EM of young and older adults in Experiment 2. The error bars represent the standard deviation.

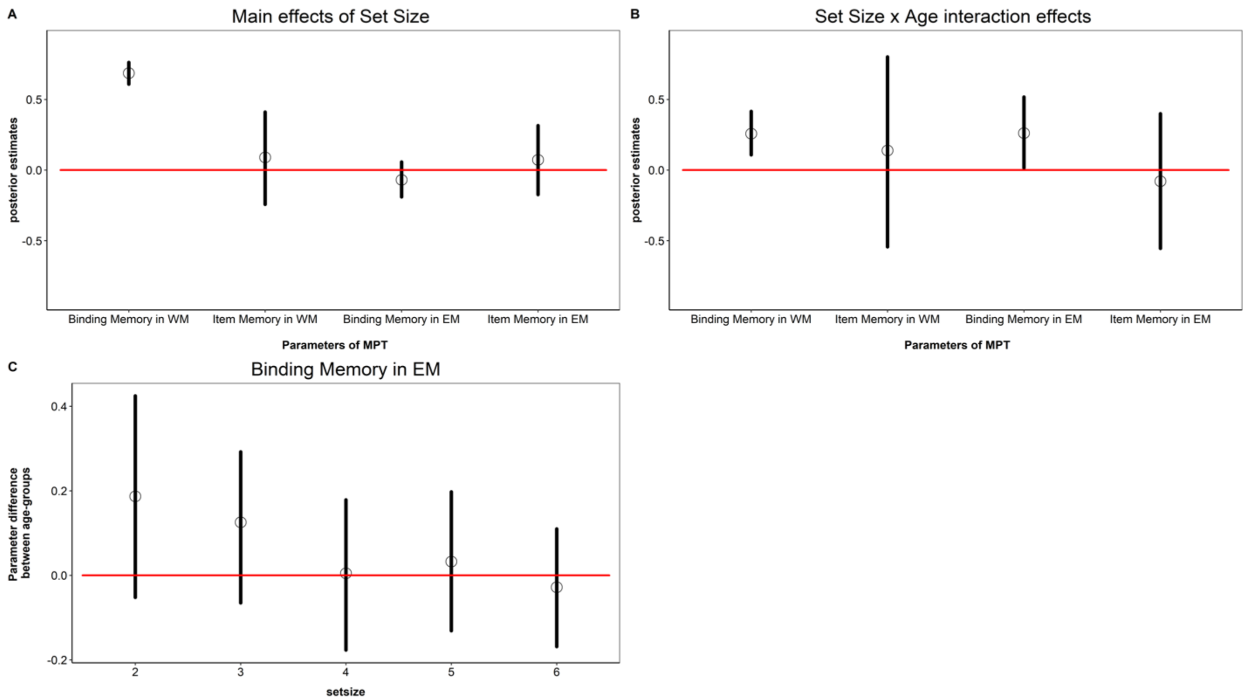

Figure 7 (A) Posterior estimates of the main effects of Set Size and (B) Posterior estimates of the interaction effects of set size with age-group of Experiment 2. Values above zero represent a stronger effect for young than for old adults. (C) Difference in posterior estimates of the parameter for binding memory in EM between the age groups per set size. Values above zero reflect an advantage for young adults. The red horizontal line characterizes the point of no evidence for an effect. The error bars represent the highest density regions.