(Only) time can tell: Age differences in false memory are magnified at longer delays

Yana Fandakova[1], Markus Werkle-Bergner[1], & Myriam C. Sander[1]

[1] Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

Correspondence concerning this article should be addressed to:

Yana Fandakova (fandakova@mpib-berlin.mpg.de) and Myriam C. Sander (sander@mpib-berlin.mpg.de) at the Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

Abstract

Older adults often report memories of past events that are partly or entirely false. To date, age differences in false memory have been examined primarily soon after the initial event. However, in real-life situations we rely on memories acquired across a lifetime. We examined associative memory in younger and older adults after 24 hours and 8 days. Age differences in memory were magnified after eight days due to a disproportionate increase in false memory rates in older adults. In both age groups, the effects of delay were modulated by memory fidelity and whether or not participants had experienced similar events potentially causing interference. Older adults were particularly vulnerable to false memory having experienced similar events, even when the initial memory was of high fidelity. We suggest that the fidelity of memory representations in concert with monitoring processes to resolve interference determine how the passage of time affects false memory.

Words: 149

Keywords: *false memory, aging, decay, interference, fidelity*

With aging, our ability to remember details about the past, such as where, when, or how specific events happened, declines (Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995). Older adults show worse memory for the binding among different features of an event accompanied by a propensity to remember associative information that is partly or entirely false (Devitt & Schacter, 2016; Dodson, 2017; Fandakova et al., 2013; Shing et al., 2008, 2009). Accordingly, memory complaints are among the most prevalent complaints of older adults (Fritsch et al., 2014; Hertzog, Hülür, Gerstof, & Pearman, 2018). Unsurprisingly, falsely remembering the past, e.g., when recalling medical instructions, or whether you locked your front door, can affect decisions and actions, and endangers independent living and quality of life in later adulthood (Farias, Mungas, & Jagust, 2005).

To date, age differences in false memory have been examined primarily at relatively short delays within 24 hours of the initial event. However, real-life decisions typically rely on memories acquired across longer periods. Does a longer interval between memory formation and retrieval differentially influence older adults' propensity for memory errors? Here, we tested the hypothesis that elders show an enhanced susceptibility to false memories after longer delays, resulting in the magnification of age differences in associative memory over time.

There are two main sources of age differences in false associative memory: Differences in the fidelity of underlying memory representations (Benjamin, 2010; Fandakova et al., 2018; Li, Naveh-Benjamin, & Lindenberger, 2005) and less efficient cognitive control processes to monitor memory retrieval (Dodson, 2017; Giovanello & Schacter, 2012; Gutchess et al., 2007; Mitchell & Johnson, 2009). The latter is particularly important when the fidelity of the underlying memory representations is low, e.g., when mnemonic evidence is scarce (Rugg, 2004) or when memories are less distinct and very similar to each other (Dulas & Duarte, 2016).

But how does the passage of time between encoding and retrieval influence memory fidelity? Decay leads to loss of detail and coherence of memory traces upon longer delays (Sadeh et al., 2014; Sekeres et al., 2016; Winocur & Moscovitch, 2011). As associative memory

fidelity decreases, we expect the likelihood for false memory to increase at longer versus shorter delays. Changes in memory over time may also critically depend on the fidelity of the initial memory representation. For example, decay may decrease the probability of high-fidelity association recall whereas an initially lower-fidelity association may become completely inaccessible over time (Habib & Nyberg, 2007; Yang et al., 2016). Evidence on age differences in forgetting over time is mixed (Elliott, Isaac, & Muhlert, 2014): Do younger and older adults actually differ in the rate with which associative representations lose fidelity over time?

In reality, of course, we rarely retrieve past events only once after a prolonged period of time. Rather, in most everyday situations we experience similar events that share cognitive representations with the original event, such as parking one's car in the same street every evening, but in a slightly different location. Such daily routines create strong representational overlap and render it difficult to recall the car's actual location when leaving the house for work in the morning (Yassa & Reagh, 2013). From an ecological point of view, the likelihood that we experience the same or similar episodes and thereby also retrieve an earlier episode, increases over time (Hardt, Nader, & Nadel, 2013). Thus, experiencing similar events induces moments of instability when new information may be partially incorporated in the original memory trace, paving the way for false memories. One way to study overlapping representations, or interference, over longer delays, is by presenting participants with additional materials that are similar to the target events (Sadeh et al., 2014). We expected that having processed similar materials would result in overlapping representations and greater false memory, especially for associations with initially low fidelity, which may be more susceptible to interference (Sadeh, Ozubko, Winocur, & Moscovitch, 2016). Resolving interference as a result of encountering similar materials requires greater engagement of monitoring processes to avoid errors (e.g., Fandakova, Lindenberger, & Shing, 2014). We expected this to be particularly challenging for older adults, resulting in more false memories (Fandakova et al., 2014). Critically, greater

representational overlap at longer delays accompanied by monitoring deficits should have an exacerbating effect on false memory, thus creating a double-jeopardy situation for older adults.

Taken together, older adults commit more false memories than younger adults do at shorter delays up to 24 hours (Devitt & Schacter, 2016). To date, it is unknown how extended delays affect age differences in false memory. Such delays are associated with loss of mnemonic fidelity through (1) decay and/or (2) increasing representational overlap via experience with similar events. While both of these mechanisms may affect younger and older adults differentially, we expected older adults to be particularly vulnerable to false associative memories for initially low-fidelity representations that become less detailed over time and therefore put high demands on monitoring processes.

Methods

Participants

Participants in this study were 31 younger adults (20–25 years) and 39 older adults (68–73 years). Sample size was chosen in line with our previous study (Fandakova et al., 2018) using a similar paradigm, demonstrating how memory fidelity influences age differences in false memory after short delays. In order to ensure that at least 20 picture–word pairs in each of the high- and low-fidelity conditions were available for the recognition test on Day 8, older participants had to achieve between 20% and 80% accuracy and younger participants between 13% and 87% accuracy on Day 1. Two younger adults and six older adults did not perform within these ranges and were excluded from further analyses. Five additional older adults were excluded for technical reasons, and two older adults dropped out before completing all assessments. Finally, two younger adults demonstrated a low overall memory performance on both Day 2 and Day 8 (defined as $p < .001$ relative to their age group) and were removed from further analyses. The final sample consisted of 27 younger adults (14 female, $M(SD)_{age}$ = 23.34(1.38) years) and 26 older adults (14 female, $M(SD)_{age}$ = 72.02(1.67) years). Participants were native German

speakers, had normal or corrected-to-normal vision, no history of psychiatric or neurological disease, and did not take psychiatric medication. Older adults were screened for cognitive impairment with the Mini-Mental State Exam (Folstein et al., 1975) and all scored above 26 points ($M(SD)_{MMSE}$ = 29.08(0.80)). The study was approved by the Ethics Committee of the Max Planck Institute for Human Development, Berlin, Germany.

Associative recognition task

We tested younger and older adults' memory for picture–word pairs after 24 hours and after 8 days of the initial associative learning (see Figure 1).
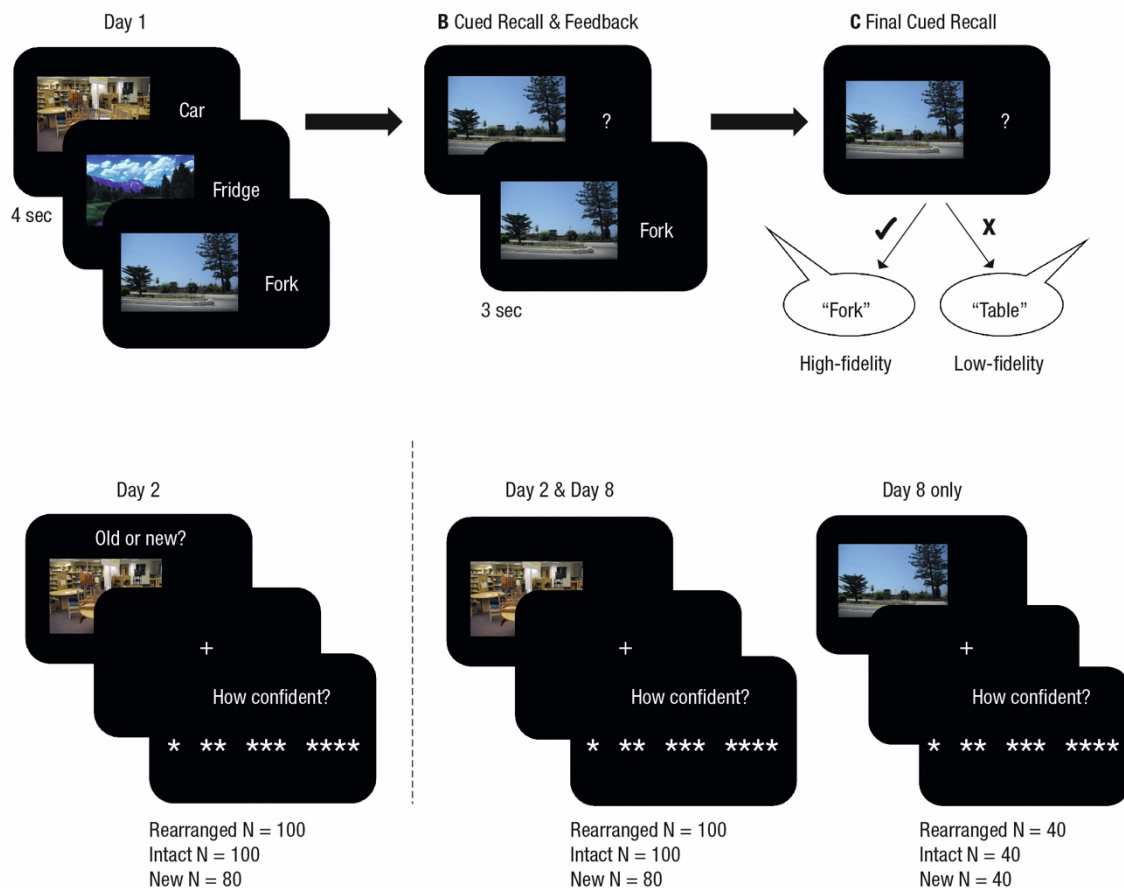

Figure 1. Experimental paradigm. Younger and older adults memorized random scene–word pairs using an imagery strategy (Day 1). 24 hours after learning (Day 2), participants completed an associative recognition task for a subset of the scene–word pairs from Day 1. On Day 8, participants returned for another associative recognition test that included pairs they had seen on Day 2 (Day 2 & Day 8) as well as pairs that were studied on Day 1, but not

presented on Day 2 (Day 8 only). In each associative recognition task, intact, rearranged, and new pairs were presented intermixed.

Day 1

Participants were first instructed to memorize random scene–word pairs using an imagery strategy. The scenes consisted of indoor and outdoor pictures and were randomly paired with concrete two-syllable nouns (for more details, see Fandakova et al., 2018). Participants were encouraged to generate integrated vivid images of the pairs and practiced using the imagery strategy prior to the main experiment. During the initial learning phase, scene–word pairs were presented for 4 seconds. Participants were instructed to remember the pair and to indicate how well they could form an integrated image of the scene–word pair on a 4-point scale. Participants then performed cued-recall blocks (one for younger adults, two for older adults). Here, scenes served as cues to verbally recall the associated word. Recall time was not constrained. Independent of recall accuracy, the correct scene–word pair was presented again for 3 seconds, fostering further learning of the pair. Following this, participants completed a final cued-recall phase without feedback. Here, scenes served as cues to recall the associated word but no feedback was provided.

As in previous studies with this paradigm (Fandakova et al., 2018; Muehlroth et al., 2019), task difficulty was adjusted between age groups to achieve recall success close to 50% in each age groups. Younger adults studied 440 pairs and older adults studied 280 pairs. In addition, younger adults completed one cued-recall block with feedback, whereas older adults completed two.

We used performance on the final cued-recall phase on Day 1 to form intact and rearranged pairs for associative recognition tests on Day 2 and Day 8. Scene–word pairs presented on Day 1 were sorted into two categories based on each individual's performance in the final cued recall on Day 1: (1) high-fidelity pairs for which participants correctly recalled the corresponding word when presented with the scene; (2) low-fidelity pairs for which participants

did not successfully recall the word when presented with the scene. For older adults, all 280

studied pairs were assigned to one of these categories. Next, 40 high-fidelity and 40 low-fidelity

pairs were set aside to be used on Day 8 as intact and rearranged pairs (20 per condition). The

remaining 200 pairs were divided equally between the intact and rearranged pair conditions on

Day 2. For younger adults, we randomly chose 100 high-fidelity and 100 low-fidelity pairs for Day

2 as well as 40 high-fidelity and 40 low-fidelity pairs for Day 8 from the larger set of studied pairs.

If younger adults' learning success did not allow this, we split the available high- and low-fidelity

pairs in half between intact and rearranged pairs for Day 2.

Our experimental manipulation on Day 1 was successful in creating variability in

associative binding success in both age groups, with $M(SD)_{younger}$ = 52.65%(18.99%), $M(SD)_{older}$

= 43.67%(16.29%), thereby allowing us to create sets of intact and rearranged pairs that

included both high- and low-quality scene–word pairs uniquely determined for each participant.

Final recall on Day 1 did not differ significantly between the age groups, $t(50)$ = 1.83, p = .07, d =

0.51.

Day 2

Approximately 24 hours after learning, participants completed an associative recognition task

with scene–word pairs from Day 1. Participants were presented with 100 intact scene–word

pairs (i.e., scene and word presented together as a pair on Day 1; consisting of parts of high-

and low-quality pairs), 100 rearranged pairs (i.e., scene and word presented as parts of different

pairs on Day 1; consisting of either high- or low-fidelity pairs), and 80 new pairs (i.e., scenes and

words had not been seen before).

Scene–word pairs were presented for 3.5 seconds, and participants indicated whether

the pair was old (i.e., an intact pair) or new (i.e., a rearranged or new pair). Participants then

rated their confidence in their decision on a 4-point scale (unlimited response time). Jittered

fixation cross periods (500-6500 ms) were presented after each memory and confidence

decision. The task was completed in five blocks of 56 trials (20 intact, 20 rearranged, 16 new pairs).

Day 8

Approximately 7 days after the associative recognition task on Day 2, participants returned for another associative recognition test. Here, they were presented with the 100 intact scene–word pairs they had seen on Day 2 as well as with 40 intact pairs (20 high- and 20 low-fidelity) that were studied on Day 1, but not presented on Day 2. In addition, participants viewed the 100 rearranged pairs presented on Day 2 along with 40 rearranged pairs (20 high- and 20 low-fidelity) they had not been seen before, but came from scenes and words they had learned on Day 1.

Finally, 40 new pairs were presented along with the 80 new pairs from Day 2. Like Day 2, scene–word pairs were presented for 3.5 seconds, followed by a confidence rating on a 4-point scale and a jittered fixation cross. Intact, rearranged, and new pairs presented on both Day 2 and Day 8 as well as those only presented on Day 8 were intermixed and presented in a different order than on Day 2. Participants were instructed to respond by saying "old" *only* to pairs that were exactly the same as on Day 1, irrespective of whether they had seen them on Day 2 as well. A practice block preceded the main task to ensure that participants understood and could perform the task. The task was completed in five blocks of 80 recognition trials each (28 intact, 28 rearranged, 16 new pairs from Day 2, and 8 completely new pairs).

Comparing pairs tested once on Day 2 or only once on Day 8 allowed us to evaluate the effects of decay on age differences in false memory. Comparing pairs presented on both Day 2 and Day 8 allowed us to evaluate the effects of interference on age differences in false memory. The alpha level for all statistical analyses was set to $\alpha = .05$ with a false discovery rate correction for multiple comparisons of post-hoc tests (indicated by p_{corr}). The Huynh-Feldt corrected degrees of freedom and significance values are reported when the sphericity assumption was not met (Huynh & Feldt, 1976).

Results

Long delays particularly increase false memories in older compared to younger adults

We first sought to determine if memory performance differentially decreased between Day 2 and Day 8 in younger and older adults. We computed recognition performance (proportion of hits, or correct "old" responses to intact pairs, minus false alarms, or incorrect "old" responses to rearranged pairs). An ANOVA on recognition performance with day (Day 2 vs. Day 8) and age group (younger vs. older adults) revealed a main effect of day, $F(1,51) = 126.21$, $p < .001$, $\eta_p^2 = .71$, indicating that memory performance decreased with longer delays in both groups (Figure 2A). We also observed a main effect of age group, $F(1,51) = 22.91$, $p < .001$, $\eta_p^2 = .31$, with overall lower memory in older adults. Notably, there was a reliable Age Group x Day interaction, $F(1,51) = 21.61$, $p < .001$, $\eta_p^2 = .30$, suggesting that longer delays had a greater negative effect on older adults' performance.
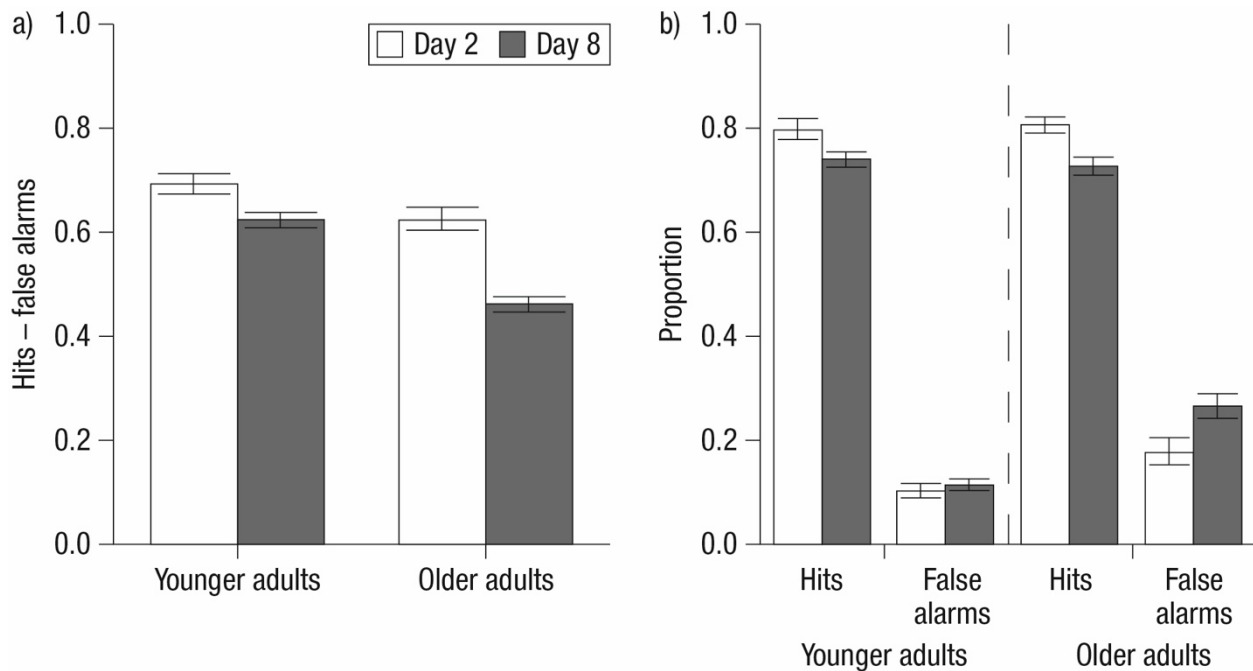

Figure 2. Associative memory on Day 2 and Day 8 for younger and older adults. (a) Overall recognition performance on Day 2 (light bars) and Day 8 (dark bars). (b) Hits and false alarms on Day 2 (light bars) and Day 8 (dark bars), separated by age group.

Did older adults' performance decline more over time due to increasing false memories, decreasing true memories, or both? To answer this question, we directly compared hits and false alarms over time. An ANOVA with memory type (hits vs. false alarms), day (Day 2 vs. Day 8), and age group (younger vs. older adults) revealed an interaction between memory type, day, and age group, $F(1,51) = 21.61$, $p < .001$, $\eta_p^2 = .30$. We followed up by performing ANOVAs with day (Day 2 vs. Day 8) and age group (younger vs. older adults) separately for hits and false alarms. A main effect of day was present for both hits, $F(1,51) = 39.64$, $p < .001$, $\eta_p^2 = .44$, and false alarms, $F(1,51) = 27.07$, $p < .001$, $\eta_p^2 = .35$, suggesting that delay resulted in decreased true and increased false associative memory. A reliable Day x Age Group interaction was present for false alarms, $F(1,51) = 16.597$, $p < .001$, $\eta_p^2 = .25$. A closer look at Figure 2b suggests that false alarms increased over time in older adults, but only minimally changed in younger adults. In contrast, hits decreased similarly in younger and older adults over time, and there was no reliable Age Group x Day interaction, $F(1,51) = 0.16$, $p = .34$, $\eta_p^2 = .02$. Together, these results indicate that longer delays between encoding and retrieval affected older adults' performance more due to greater increases in false memories over time.

Magnification of age differences in false memory over time with representational overlap

Next, we examined age differences in false associative memory at shorter and longer delays depending on memory fidelity.

Upon a shorter delay (Day 2), an ANOVA on false alarms to rearranged pairs with memory fidelity (high vs. low) and age group (younger vs. older adults) demonstrated a main effect of fidelity, $F(1,51) = 79.14$, $p < .001$, $\eta_p^2 = .61$ such that both groups committed more false alarms for low- than for high-fidelity pairs (Figure 3a). There was also a main effect of age group, $F(1,51) = 6.99$, $p = .01$, $\eta_p^2 = .12$, indicating that older adults were overall more likely to falsely claim that a rearranged pair has been seen before. The Fidelity x Age Group interaction was not significant, $p = .77$.

Next, we examined how false memory changed over time. The dependent variable in this analysis was change in false alarms on Day 8 relative to Day 2 for the corresponding fidelity type (see Figure 3b). Here, we were particularly interested in determining whether decay or representational overlap had a differential effect on false memory in younger and older adults. An ANOVA with memory fidelity (low vs. high), condition (Day 2 & Day 8 vs. Day 8 only) and age group (younger vs. older adults) showed a main effect of age group, $F(1,51) = 9.39$, $p = .003$, $\eta_p^2 = .16$, consistent with the overall greater increase in false alarms in older adults reported above. There was also a main effect of condition, $F(1,51) = 96.02$, $p < .001$, $\eta_p^2 = .65$, indicating that false memory increased more for pairs presented on Day 2 and Day 8 than those tested on Day 8 only. We also found a main effect of fidelity, $F(1,51) = 23.33$, $p < 0.001$, $\eta_p^2 = .31$, and a three-way interaction between age group, condition, and fidelity, $F(1,51) = 5.89$, $p = .019$, $\eta_p^2 = .10$. To unpack these effects, we examined how age and memory fidelity influenced change in false memory within each condition (Day 2 & Day 8 vs. Day 8 only).

Decay (Day 8 only). An ANOVA with fidelity (high vs. low) and age group (younger vs. older adults) on change in false alarms relative to Day 2 revealed a main effect of fidelity, $F(1,51) = 36.60$, $p < .001$, $\eta_p^2 = .42$. False alarms to high-fidelity pairs were similar or even more frequent when tested on Day 8 days than on Day 2. In contrast, false alarms to low-quality pairs were actually more rare when tested on Day 8 relative to Day 2 in both younger and older adults. There was no Fidelity x Age Group interaction, $p = .97$. The main effect of age group was reliable, $p < .001$, consistent with the overall ANOVA results reported above.

Representational overlap (Day 2 & Day 8). An ANOVA with fidelity (high vs. low) and age group (younger vs. older adults) revealed a reliable Age Group x Fidelity interaction, $F(1,51) = 9.52$, $p = .003$, $\eta_p^2 = .16$. Post-hoc tests revealed that younger and older adults showed similar increases in low-fidelity false alarms between Day 2 and Day 8 (Figure 3b), $p_{corr} = .54$, $d = .17$. In contrast, false memory of high-quality pairs increased more in older than in younger adults, p_{corr}

< .001, *d* = 1.68. The age group effect was again reliable, *p* < .001, consistent with the ANOVA

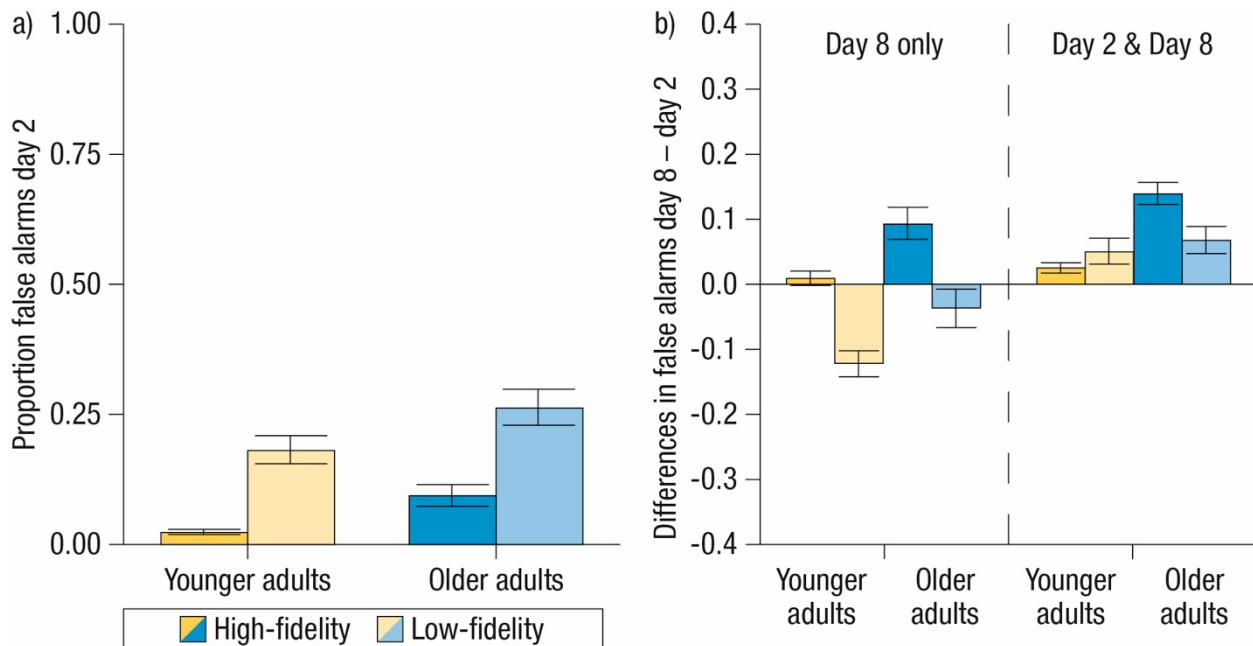results. The main effect of fidelity was not significant, *p* = .147.

Figure 3. False alarms on Day 2 and Day 8, separated by memory fidelity. (a) Proportion of false alarms for high-fidelity (darker colors) and low-fidelity (lighter colors) pairs on Day 2 across younger and older adults. (b) Differences in false alarms on Day 8 relative to Day 2 for high-fidelity (darker colors) and low-fidelity pairs (lighter colors) tested on Day 8 only (left) and those tested on Day 2 & Day 8 (right).

In the representational overlap condition, participants encountered rearranged pairs on

both Day 2 and Day 8, and made old/new decisions on these pairs on both days. Here, Day 2

recognition success may have affected the likelihood to incorrectly endorse rearranged pairs on

Day 8. We performed control analyses to compare the proportion of false alarms on Day 2 that

remained false alarms on Day 8 and the proportion of correct rejections on Day 2 that

transitioned into false alarms on Day 8. An ANOVA with transition (false–false vs. correct–false)

and age group (younger vs. older adults) demonstrated a significant effect of transition, $F(1,51)$

= 22.816, *p* < .001, η_p^2 = .31, along with a reliable Age Group x Transition interaction, $F(1,102)$ =

9.23, *p* = .004, η_p^2 = .15. False alarms on Day 2 were generally more likely to remain false

alarms in younger, $M(SD)_{younger}$ = .61(.28), and older adults, $M(SD)_{older}$ = .55(.17), p_{corr} = .37, d = 0.22. Fewer correct rejections on Day 2 transitioned to false alarms on Day 8 for younger, $M(SD)_{younger}$ = .09(.05), than for older adults, $M(SD)_{older}$ = .24(.10), p_{corr} < .001, d = 1.90.

What might contribute to this higher transition likelihood in older adults? One possibility is that the transitioning pairs were less confident correct rejections on Day 2 in the first place (i.e., guesses), making the original memory trace more susceptible to integration with the presented rearranged information on Day 2. If so, items transitioning from correct to false between Day 2 and Day 8 would be endorsed with lower confidence on Day 2. To test this, we examined the level of confidence for correct rejections on Day 2 based on whether they subsequently remained correct rejections or transitioned to false alarms on Day 8. An ANOVA on Day 2 confidence judgements by transition (correct–correct vs correct–false) and age group (younger vs. older adults) revealed a main effect of transition, $F(1,51)$ = 20.59, p < .001, η_p^2 = .29 along with an Age Group x Transition interaction, $F(1,51)$ = 13.15, p = .001, η_p^2 = .21. Differences between younger and older adults were minimal for correct–correct transitions, p_{corr} = .05, d = 0.57, with older adults, $M(SD)_{older}$ = 3.32(0.45), showing slightly *lower* confidence than younger adults, $M(SD)_{younger}$ = 3.53(0.27). By contrast, older adults reported *higher* confidence on Day 2 for correct rejections that later became false alarms, $M(SD)_{older}$ = 3.10(0.55) than did younger adults $M(SD)_{younger}$ = 2.79(0.67), p_{corr} = .01, d = .51. Together, these exploratory analyses suggest that correctly rejected pairs that were later falsely remembered were associated with lower confidence on Day 2 in younger adults, whereas confidence was not predictive of the future fate of the memory trace in older adults.

Discussion

This study provides novel evidence that age differences in false memory persist and become magnified over time such that older adults demonstrate an even greater susceptibility to false associative memory when tested after longer delays. Compared to younger adults, older adults' associative memory declined more 8 days after learning than 24 hours after learning.

This finding is consistent with the accelerated forgetting of associative information in older adults shown by Mary, Schreiner, and Peigneux (2013) and contributes to an ongoing debate whether aging is associated with accelerated long-term forgetting (e.g., Elliott et al., 2014). Directly comparison of true and false memory revealed that the magnification of age differences over time was primarily due to older adults' propensity to wrongly remember unstudied associations. This highlights the importance of considering true and false memory separately when examining associative memory at longer delays.

Whereas previous research has suggested that decay and representational overlap are the major sources of changes in the quality of mnemonic information over time (Hardt et al., 2013; Sadeh et al., 2014), evidence how these factors modulate age differences in false memory is scarce. This study demonstrates that false memories depend on the fidelity of the associative memory representation formed at encoding and are particularly frequent in older adults when interference due to overlapping representations has to be resolved.

Decay effects on false memory depend on initial memory fidelity

Associative memory fidelity critically influences older adults' propensity for false memory when tested 24 hours after encoding (Fandakova et al., 2018). Here, we extended the delay to 8 days. For high-fidelity associations, false alarms remained stable or increased after 8 days relative to 24 hours. In contrast, errors for low-fidelity associations actually *decreased* over time in younger and older adults. These results underscore that decay effects on false memory are not uniform, but depend on the fidelity of the initially formed representation. These results resonate with research using the Deese-Roediger-McDermott paradigm in younger adults suggesting that false memory may be more stable than true memory over time (e.g., Seamon et al., 2002).

Details, especially those not central to an event, are forgotten over time in younger adults (Sekeres et al., 2016). Therefore, we speculated that memory fidelity would decrease over time. As a result, initially high-quality associative memories would become more similar to low-quality

memories, whereas initially low-quality memories would decline, potentially rendering those pairs inaccessible during retrieval (e.g., Habib & Nyberg, 2007). False memories are unlikely at both ends of the continuum of memory fidelity: They are highly unlikely for well-learned associations (Fandakova et al., 2018; Rotello et al., 2000) but also when the memory trace is weak or lacking mnemonic evidence does not suffice. In the case of completely new information as an extreme example, novelty can even facilitate responses and help to avoid memory errors (e.g., Dobbins, Kroll, Yonelinas, & Liu, 1998; Tulving & Kroll, 1995). Our finding of decreasing false memory for low-fidelity pair representations indicate that these pairs further lose fidelity across time, reducing the probability for false claims that recombined information was old. In line with this interpretation, low-fidelity hits showed pronounced decreases over time in both age groups (see Supplementary Materials). Together, our results suggest that the effect of fidelity on memory error likelihood may follow an inverted U-shape with a maximum when some familiarity threshold is reached (Figure 4). Accordingly, the effect of decay with time simply results in a shift of associative memories across the inverted U-shape relating false alarms and fidelity.
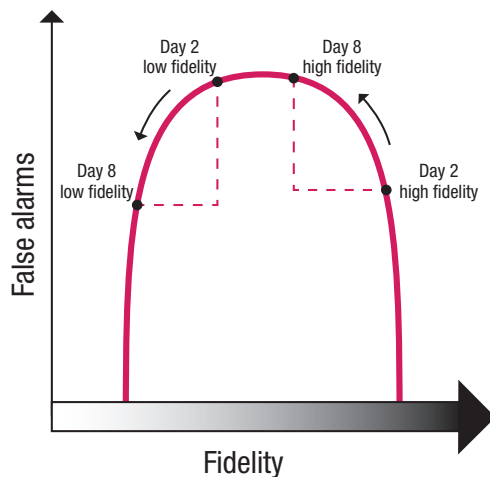

Figure 4. Schema of how longer versus shorter delays may affect the level of fidelity for a given memory representation and thereby modulate false associative memory.

Effects of representational overlap on false memory differ between younger and older adults

Memory fidelity does not only passively decay over time. Rather, mnemonic information is modified when we experience similar events, resulting in overlapping representations and increased interference. In this study, we simulated the daily experience of similar events by testing rearrangements of some of the scene–word pairs at two time points. We expected this manipulation to increase representational overlap between the initial representation and the rearranged pairs, thereby also increasing potential interference at retrieval.

The pattern observed in younger adults is consistent with the idea that presenting similar information in the form of rearranged pairs on Day 2 has stronger effects on low-fidelity memories that are more susceptible to new erroneous information during the associative test on Day 2. In contrast, the minimal increase in false memory for high-fidelity pairs suggests that a well-bound representation is resilient to representational similarity. This finding is in line with suggestions by Sadeh and colleagues (2014) that interference effects depend on the quality of the underlying memory trace.

Previous research suggested that age differences in monitoring processes at retrieval contribute to differences in false memory after short delays (e.g., Devitt & Schacter, 2016). Here, presenting rearranged pairs on Day 2 was expected to induce incorporation of erroneous information requiring more monitoring for successful interference resolution. In line with previous work (Duarte & Dulas, 2016; Fandakova et al., 2018), we expected older adults to have more difficulties in successfully engaging monitoring processes, whereas younger adults' well-bound associative representations would be protected from interference. However, older adults showed a disproportionate increase in false alarms for high-fidelity memories. This suggests that deficits in retrieval monitoring processes contribute to age differences in associative memory, and that impaired memory fidelity is unlikely to be the only source of elevated false alarms in older adults. The exploratory analysis of confidence ratings lends further support for this. Younger adults'

correct rejections on Day 2 that transitioned to false alarms on Day 8 were associated with lower subjective confidence, potentially reflecting the relative lack of mnemonic evidence compared to pairs that were correctly rejected on both days. No such effects were observed in older adults where confidence judgments did not differentiate between correct rejections Day 2 that remained, or transitioned to false alarms on Day 8. This result is consistent with findings that age-related monitoring deficits at retrieval are often associated with miscalibrated confidence judgments in older adults (Dodson et al., 2007; Fandakova et al., 2013; Shing et al., 2009). Interestingly, confidence may initiate further information seeking in decision-making (e.g., Desender, Boldt, & Yeung, 2018) or during self-guided learning (e.g., Bjork, Dunlosky, & Kornell, 2013). Thus, Day 2 confidence judgments may potentially tag information for retention in its modified form.

Taken together, our results suggest that testing associative memory after longer delays creates a double jeopardy for older adults: first, they are more likely to find themselves in situations with greater similarities between episodes that increase interference resolution demands. Furthermore, age-related declines in monitoring processes impair older adults' effective interference resolution, thereby raising future false memory likelihood. Older adults are thus particularly vulnerable to misremembering episodes over time. Our findings highlight the necessity to find effective ways to increase memory fidelity and to structure older adults' environment in ways that decrease demands on monitoring.

References

Benjamin, A. S. (2010). Representational explanations of "process" dissociations in recognition: The DRYAD theory of aging and memory judgments. *Psychological Review, 117*(4), 1055–1079. https://doi.org/10.1037/a0020810

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science, 29*(5), 761–778. https://doi.org/10.1177/0956797617744771

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia, 91*, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Dobbins, I. G., Kroll, N. E. A., Yonelinas, A. P., & Liu, Q. (1998). Distinctiveness in recognition and free recall: The role of recollection in the rejection of the familiar. *Journal of Memory and Language, 38*(4), 381–400. https://doi.org/10.1006/jmla.1997.2554

Dodson, C. S. (2017). Aging and memory. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed., pp. 403–421). Oxford: Academic Press.

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging, 22*(1), 122–133. https://doi.org/10.1037/0882-7974.22.1.122

Dulas, M. R., & Duarte, A. (2016). Age-related changes in overcoming proactive interference in associative memory: The role of PFC-mediated executive control processes at retrieval. *NeuroImage, 132*, 116–128. https://doi.org/10.1016/j.neuroimage.2016.02.017

Elliott, G., Isaac, C. L., & Muhlert, N. (2014). Measuring forgetting: A critical review of accelerated long-term forgetting studies. *Cortex, 54*, 16–32. https://doi.org/10.1016/j.cortex.2014.02.001

Fandakova, Y., Lindenberger, U., & Shing, Y. L. (2014). Deficits in process-specific prefrontal and hippocampal activations contribute to adult age differences in episodic memory interference. *Cerebral Cortex, 24*(7), 1832–1844. https://doi.org/10.1093/cercor/bht034

Fandakova, Y., Sander, M. C., Grandy, T. H., Cabeza, R., Werkle-Bergner, M., & Shing, Y. L. (2018). Age differences in false memory: The importance of retrieval monitoring processes and their modulation by memory quality. *Psychology and Aging, 33*(1), 119–133. https://doi.org/10.1037/pag0000212

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013). High-confidence memory errors in old age: The roles of monitoring and binding processes. *Memory, 21*(6), 732–750. https://doi.org/10.1080/09658211.2012.756038

Farias, S. T., Mungas, D., & Jagust, W. (2005). Degree of discrepancy between self and other-reported everyday functioning by cognitive status: Dementia, mild cognitive impairment, and healthy elders. *International Journal of Geriatric Psychiatry, 20*(9), 827-834. https://doi.org/10.1002/gps.1367

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*, 189–198.

Fritsch, T., McClendon, M. J., Wallendal, M. S., Hyde, T. F., & Larsen, J. D. (2014). Prevalence and cognitive bases of subjective memory complaints in older adults: Evidence from a community sample. *Journal of Neurodegenerative Diseases, 2014*, 176843. 10.1155/2014/176843

Giovanello, K. S., & Schacter, D. L. (2012). Reduced specificity of hippocampal and posterior ventrolateral prefrontal activity during relational retrieval in normal aging. *Journal of Cognitive Neuroscience, 24*(1), 159–170. https://doi.org/10.1162/Jocn_a_00113

Gutchess, A. H., Hebrank, A., Sutton, B. P., Leshikar, E., Chee, M. W. L., Tan, J. C., ... Park, D. C. (2007). Contextual interference in recognition memory with age. *Neuroimage, 35*(3), 1338–1347. https://doi.org/10.1016/j.neuroimage.2007.01.043

Habib, R., & Nyberg, L. (2008). Neural correlates of availability and accessibility in memory. *Cerebral Cortex, 18*(7), 1720–1726. https://doi.org/10.1093/cercor/bhm201

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: The role of active forgetting in memory. *Trends in Cognitive Sciences, 17*(3), 111–120. https://doi.org/10.1016/j.tics.2013.01.001

Hertzog, C., Hulur, G., Gerstorf, D., & Pearman, A. M. (2018). Is subjective memory change in old age based on accurate monitoring of age-related memory change? Evidence from two longitudinal studies. *Psychology and Aging, 33*(2), 273–287. https://doi.org/10.1037/pag0000232

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1*(1), 69–82. https://doi.org/10.3102/10769986001001069

Li, S.-C., Naveh-Benjamin, M., & Lindenberger, U. (2005). Aging neuromodulation impairs associative binding: A neurocomputational account. *Psychological Science, 16*(6), 445–450. https://doi.org/10.1111/j.0956-7976.2005.01555.x

Mary, A., Schreiner, S., & Peigneux, P. (2013). Accelerated long-term forgetting in aging and intra-sleep awakenings. *Frontiers in Psychology, 4*, 750. https://doi.org/10.3389/fpsyg.2013.00750

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin, 135*(4), 638–677. https://doi.org/10.1037/a0015849

Muehlroth, B. E., Sander, M. C., Fandakova, Y., Grandy, T. H., Rasch, B., Shing, Y. L., & Werkle-Bergner, M. (2019). Precise slow oscillation-spindle coupling promotes memory consolidation in younger and older adults. *Scientific Reports, 9*, 1940. https://doi.org/10.1038/s41598-018-36557-z

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

Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language, 43*(1), 67–88. https://doi.org/10.1006/jmla.1999.2701

Rugg, M. D. (2004). Retrieval processing in human memory: Electrophysiological and fMRI evidence. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 727–737). Cambridge, MA: MIT Press.

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2014). How we forget may depend on how we remember. *Trends in Cognitive Sciences, 18*(1), 26–36. https://doi.org/10.1016/j.tics.2013.10.008

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2016). Forgetting patterns differentiate between two forms of memory representation. *Psychological Science, 27*(6), 810–820. https://doi.org/10.1177/0956797616638307

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition, 30*(7), 1054–1064. 10.3758/BF03194323

Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., & Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: Differential rates of forgetting for detail types in episodic memory. *Learning & Memory, 23*(2), 72–82. https://doi.org/10.1101/lm.039057

Shing, Y. L., Werkle-Bergner, M., Li, S. C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A life-span dissociation. *Journal of Experimental Psychology: General, 137*(3), 495–513. https://doi.org/10.1037/0096-3445.137.3.495

Shing, Y. L., Werkle-Bergner, M., Li, S. C., & Lindenberger, U. (2009). Committing memory errors with high confidence: Older adults do but children don't. *Memory, 17*(2), 169–179. https://doi.org/10.1080/09658210802190596

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. https://doi.org/10.1037/0882-7974.10.4.527

Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term-memory encoding. *Psychonomic Bulletin & Review, 2*(3), 387–390. https://doi.org/10.3758/Bf03210977

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society, 17*(5), 766–780. https://doi.org/10.1017/S1355617711000683

Yang, J. J., Zhan, L. X., Wang, Y. Y., Du, X. Y., Zhou, W. X., Ning, X. L., ... Moscovitch, M. (2016). Effects of learning experience on forgetting rates of item and associative memories. *Learning & Memory, 23*(7), 365–378. https://doi.org/10.1101/lm.041210.115

Yassa, M. A., & Reagh, Z. M. (2013). Competitive trace theory: a role for the hippocampus in contextual interference during retrieval. *Frontiers in Behavioral Neuroscience, 7:* 107. https://doi.org/10.3389/fnbeh.2013.00107

(Only) time can tell: Age differences in false memory are magnified at longer delays

Supplementary Materials

Results: Changes in true memory over time in younger and older adults

On Day 2, an ANOVA on hits with memory fidelity (high vs. low) and age group (younger vs. older adults) demonstrated a main effect of fidelity, $F(1,51) = 195.58$, $p < .001$, $\eta_p^2 = .79$, suggesting that both age groups demonstrated higher memory for high- than for low-quality pairs (Supplementary Figure 1A). There were no main effects of age group, $p = .75$, nor Fidelity x Age Group interactions, $p = .66$.

How did true memory change over time? An ANOVA on change in hits on Day 8 relative to Day 2 with fidelity (low vs. high), condition (Day 2 & Day 8 vs. Day 8 only), and age group (younger vs. older adults) revealed a main effect of condition, $F(1,51) = 274.23$, $p < .001$, $\eta_p^2 = 0.84$ along with a Fidelity x Condition interaction $F(1,51) = 69.004$, $p < .001$, $\eta_p^2 = .58$. Presentation of the same intact pairs on Day 2 resulted in preservation of performance in both age groups (Supplementary Figure 1b), and this effect was more pronounced for high-quality pairs. Additionally, there was a main effect of quality, $F(1, 51) = 21.87$, $p < .001$, $\eta_p^2 = 0.30$. Consistent with the age-uniform change in true memory over time, we found no effects of age group or interactions involving age group (all $p > .05$).
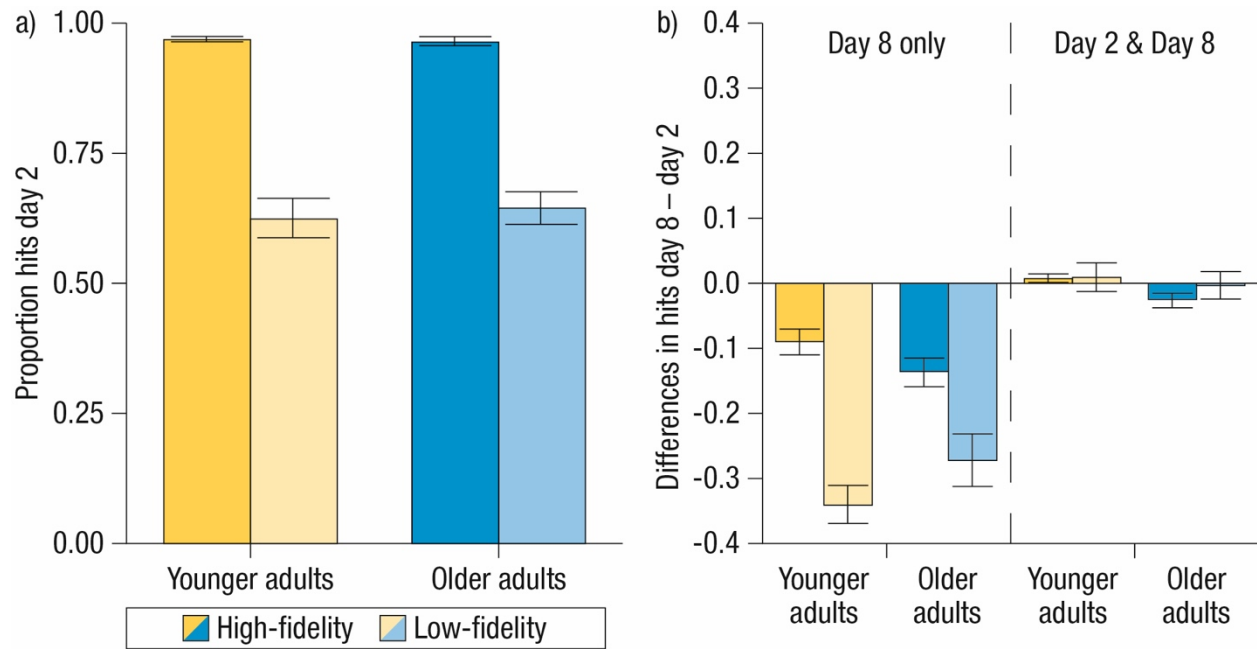
Supplementary Figure 1. Hits on Day 2 and Day 8, separated by memory fidelity. (a) Proportion of hits for high-fidelity (darker colors) and low-fidelity (lighter colors) pairs on Day 2 across younger and older adults. (b) Differences in hits on Day 8 relative to Day 2 for high-fidelity (darker colors) and low-fidelity pairs (lighter colors) tested on Day 8 only (left) and those tested on Day 2 & Day 8 (right).