Article:

Mao Chao, C. orcid.org/0000-0002-4087-6176, Xu, C., Loaiza, V. orcid.org/0000-0002-5000-7089 et al. (1 more author) (2024) Are latent working memory items retrieved from long-term memory? Quarterly Journal of Experimental Psychology, 77 (8). pp. 1703-1726. ISSN 1747-0218

https://doi.org/10.1177/17470218231217723

Are latent working memory items retrieved from long-term memory?

Chang Mao Chao*[1], Chenlingxi Xu[1], Vanessa Loaiza[2] & Nathan S. Rose[1]

[1]University of Notre Dame, [2]University of Essex

*Address Correspondence to:

Dr. Chang Mao Chao

390 Corbett Family Hall

University of Notre Dame

Notre Dame, IN 46556

cchao3@nd.edu

Abstract

Switching one's focus of attention between to-be-remembered information in working memory (WM) is critical for cognition, but the mechanisms by which this is accomplished are unclear. Some models suggest that passively retaining "latent" information outside of focal attention and returning it to the focus involves episodic long-term memory (LTM) retrieval processes even for delays of only a few seconds. We tested this hypothesis by examining performance on both a two-item, double-retrocue WM task (that oriented participants' attention to the item that would be tested first and then to the item tested second on each trial) and subsequent LTM tests for the items from the initial WM task. We compared performance on these tests between older adults (a population with LTM deficits) and young adults with either full (Experiment 1) or divided (Experiment 2) attention during the WM delay periods. Retrocueing, aging, and divided attention all had significant effects on WM performance, but did not interact with or systematically affect subsequent LTM performance for item, location, or associative memory judgments made with either high or low confidence. These dissociations between WM and LTM suggest that retaining and reactivating an item outside of focal attention on this two-item, double-retrocue WM paradigm, which has shown neuroimaging, neurostimulation, and neurocomputational modeling evidence for latent WM, does not involve LTM retrieval processes; rather, the results are consistent with the Dynamic Processing Model of WM (Rose, 2020, *Current Directions in Psychological Science*).

In everyday life, we often encounter situations where we must remember information only briefly in working memory (WM) and then possibly retrieve it later on from episodic long-term memory (LTM). For example, using two-factor authentication to access one's account often requires a 6-digit code to be sent via text message to verify the username and password. The code may be maintained temporarily in WM until it can be entered and authenticated. If attention is temporarily drawn to processing some other information (e.g., another unrelated incoming text message), the code is no longer in one's focus of attention to be actively rehearsed or retained. In this case, where does the "latent" memory of the code go? How is it represented in the mind and brain? Without having to reread the text message, how can it be brought back to mind to enter and authenticate the code? The current research aims to address this question.

A great deal of research related to this issue has focused on the intersection between attention, WM, and LTM (Baddeley, 2012; Oberauer, Lewandowsky, Awh, Brown, Conway, Cowan, & Ward, 2018). Some models propose that when attention is switched away from actively maintaining an item, returning it back into the focus of attention involves retrieving it from LTM (Atkinson & Shiffrin, 1968; Cowan, 2001; McCabe, 2008; McElree, 2006; Oberauer, 2002; Unsworth & Engle, 2007). Indeed, a recent account suggests that such "latent" memories are not “in WM” per se -- they must be retained in, and retrieved from, LTM (Foster, Vogel, & Awh, in press). Conversely, an alternative activity-silent WM account proposed by the Synaptic Theory of WM suggests that latent items may still be retained in WM via short-term synaptic plasticity mechanisms (Mongillo, Barak, & Tsodyks, 2008; Stokes, 2015; Trübutschek, Marti, Ojeda, King, Mi, Tsodyks, & Dehaene, 2017), and thus reactivation of latent items does not require retrieval from LTM. The current study aimed to distinguish between these different accounts about whether and how retrieval from LTM is involved in maintaining information outside of focal attention in WM.

Investigating the role of attention in WM using the retrocue paradigm

One method that has been used to study the roles of attention and LTM in WM is the retrocue paradigm. In these WM tasks, a retrospective-attention-cue orients participants to prioritize the maintenance of one or more items in WM. Therefore, retrocued items are thought to be held in a higher “state of activation” in the focus of attention than the other deprioritized items, and this typically provides a benefit to the retrocued items in terms of memory accuracy, precision, or response times (for review, see Souza & Oberauer, 2016). For example, Souza, Rerko, & Oberauer (2014) found retrocueing benefits on a WM task in which participants were instructed to memorize an array of colors followed by either a retrocue or a no-cue condition. In the retrocue condition, an arrow pointed to the to-be-remembered item that was subsequently tested by a probe color. Participants indicated whether or not the probe color matched the retrocued color being held in the focus of attention in WM. The results showed that retrocued items were better recognized than uncued items, which suggests that the retrocues directed participants’ attention to the cued items to prioritize their maintenance and representation in WM.

In single-retrocue paradigms (and the vast majority of other WM paradigms), the retrocues always indicate the items that are to-be-attended and retained in focal attention throughout the trial. However, the double-retrocue paradigm is particularly useful for characterizing the role of LTM in WM because it helps to de-confound the role of internally directed attention from WM retention (Lewis-Peacock, Drysdale, Oberauer & Postle, 2012; LaRocque, Lewis-Peacock & Postle, 2014; Rose, LaRocque, Riggall, Gosseries, Starrett, Meyering, & Postle, 2016). In the double-retrocue paradigm , after the participant switches their attention away from the uncued item(s) and is tested on the first cued item, a second retrocue indicates which item is to be tested next on the trial. Thus, the participant may be required to switch their attention back to the initially uncued item(s) to reactivate what was dropped from focal attention. Several models of memory assume that this reactivation involves retrieving the items using episodic retrieval processes as in tests of LTM (Atkinson & Shiffrin, 1968; Cowan, 2008; McCabe, 2008; Unsworth & Engle, 2007).[1]

---

1 The terminology used in the literature to describe these concepts varies across researchers and has evolved over time. We previously used the terms Attended and Unattended items to refer to the cued and uncued items, so we continue to use those terms for consistency here. We also connect unattended, potentially relevant, but deprioritized items as latent or passively-retained items to distinguish them from both items that are actively-retained in the focus of attention, and uncued items that are no longer relevant on the trial that can be dropped, removed, or deleted from WM. We hope that clarification with this terminology will help the field move forward to clearly convey related and distinct concepts.

To test this assumption, LaRocque, Eichenbaum, Starrett, Rose, Emrich, and Postle (2015, Experiments 2 and 3) administered a double-retrocue WM task followed by a surprise, subsequent LTM test of the memoranda from the initial WM task according to whether they were cued successively (i.e., consistently retained) or initially uncued but cued thereafter (i.e., dropped but then reactivated). The design followed the reasoning that, if retrieval from LTM is required to reactivate information that was dropped from focal attention, then subsequent LTM should be greater for these items compared to those that had been consistently retained in focal attention. In the beginning of each trial, two images of common, nameable objects were presented to healthy young adult participants and then, following a delay period, a first cue pointed to the image that was to be tested first. Participants thereafter saw a probe image and responded as to whether it was a match or nonmatch to the cued image. Then a second cue and a second probe were presented for participants to make a match/non-match response to the second cued stimulus. Following all of the trials of the WM task, participants took a surprise subsequent recognition LTM test in which all of the to-be-remembered images from the WM task and an equal number of new images were presented, one at a time, and the participants indicated whether or not each image had been presented during the WM task. The items were categorized into four conditions based on how they were initially held in different states of prioritization on the WM task: A-A (attended 1st & attended 2nd), A-U (attended 1st & unattended 2nd), U-A (unattended 1st & attended 2nd), and U-U (unattended 1st & unattended 2nd).

Subsequent LTM was compared for these items to test the hypothesis that episodic LTM processes were involved in reactivating items held in WM that were dropped from focal attention. Specifically, the U-A condition was considered as the item that was dropped from the focus of attention, but was subsequently reactivated. If LTM was engaged in this process, then performance should be better for the U-A condition than the A-U condition. However, the results showed that subsequent LTM of items from the initial WM task was similar between the U-A and A-U conditions. Based on this result, the authors concluded that recovering latent WM items did not involve LTM. However, there may be several other reasons for LaRocque and colleagues' findings: covert rehearsal was not prevented during the delay periods; participants did not know that their memory for the items from the initial WM task would be tested later (i.e., incidental, not intentional, encoding for the LTM test); and the LTM test did not assess memory for different types of details that may accompany retrieval (e.g., associated context bindings, confidence). As explained further on, the current study addresses these potential caveats and related issues.

LaRocque and colleagues' findings are consistent with some observations from neural recordings and theories based on neurocomputational models, which are helpful to determine whether items that are dropped from continuous maintenance in focal attention are represented in, and retrieved from, either WM or LTM. Maintaining information in WM has long been considered to rely on sustained, active neural representations of to-be-remembered information based on a variety of neuroscience data (e.g., Constantinidis, Funahashi, Lee, Murray, Qi, Wang & Arnsten, 2018; Fuster & Alexander, 1971; for reviews, see Stokes, 2015, and Rose, 2020). However, many double-retrocue studies show that active neural representations of uncued items return to baseline during WM delay periods, suggesting that the items are not continuously maintained in a sustained, active manner (Lewis-Peacock et al., 2012; Rose et al., 2016). In these double-retrocue paradigms, both items are decodable during the stimulus presentation period. After the first retrocue, only the cued item could be decoded during the post-cue delay period. The neural representation of the uncued items dropped to the baseline level of representation as if it were forgotten (i.e., it became indistinguishable from the amount of neural evidence for the category that was absent on that trial). Importantly, when this latent item was subsequently cued later in the trial, participants could rapidly and accurately switch to focusing attention on the item and there was a corresponding return in neural decoding (Lewis-Peacock et al., 2012; Rose et al., 2016).

The "activity-silent" short-term retention mechanisms proposed by the Synaptic Theory of WM can provide an account of the passive retention of latent items (Mongillo et al., 2008; Stokes, 2015; Trübutschek et al., 2017). This activity-silent WM account posits that latent items can be represented and briefly retained via short-term synaptic plasticity mechanisms that can modulate synaptic weights in a rapid, transient manner. The synaptic weights are represented by the influx of calcium concentration in the presynaptic terminal that depolarizes the membrane potential to be closer to threshold, thereby potentiating the cell to easily fire again if presynaptic input returns. This influx of calcium concentration is short-lived, from hundreds to thousands of milliseconds, so the synaptic weights can briefly code for stimulus specific information and then these weights are quickly cleared when the information is no longer relevant. That is, without a return of presynaptic activity, the transient synaptic weights are cleared to an un-potentiated, baseline state. These short-term plasticity mechanisms are distinct from the long-term potentiation mechanisms that involve protein synthesis required for axon and dendrite sprouting that provide the basis for LTM representation (Mongillo et al., 2008; Stokes, 2015; Trübutschek et al., 2017). Thus, latent items may be retained in an activity-silent or hidden state. This model also suggests that information in an activity-silent state can be reactivated by non-specific input to the network. Consistent with this theory, Rose et al. (2016) demonstrated that single pulses of transcranial magnetic stimulation (TMS) could reactivate latent WM items while they were still relevant on the trial, but TMS had no effect on items that were either actively retained or items that were no longer relevant on the trial (for replications and extensions, see Fulvio & Postle, 2020; Wolff, Jochim, Akyürek, & Stokes, 2017).

Is activity-silent WM just LTM?

An alternative explanation to the activity-silent WM account appeals to the involvement of LTM in WM tasks, as explained previously (Foster et al., in press; LaRocque, Lewis-Peacock & Postle, 2014; Rose, 2020). That is, the neural activity of latent items may drop to baseline because they are no longer retained 'in WM' per se. The representations of these items may be in the beginning stages of more long-lasting synaptic weight modification (i.e., long-term potentiation processes that underlie LTM consolidation). Thus, the latent WM items may be retrieved when these items are subsequently cued via LTM retrieval processes. A developing literature suggests that latent WM may be better conceptualized as LTM (Buschman, Siegel, Roy, & Miller, 2011; Emrich, Riggall, Larocque, & Postle, 2013; Foster et al, in press; Rose, 2020). For example, Foster et al. recently posited that an online-offline model of memory can account for the distinction between WM and LTM. According to this account, a limited set of items are able to be maintained online in WM, whereas items that are not in the current focus of attention are stored offline in LTM. This is consistent with several prominent WM models: Regardless of whether they view WM either as an active subset of LTM (Cowan, 2008; Oberauer, 2009) or completely distinct from LTM (Baddeley, 1986; Barrouillet & Camos, 2015), many models assume that information that is not “in WM” necessarily must be retrieved from LTM.

As alluded to previously, considerable cognitive research may be seen to provide support for this logic. For example, some work has shown that retrieval from LTM is greater for items initially studied during complex span tasks than simple span tasks, known as the McCabe effect (e.g., Loaiza & McCabe, 2012; McCabe, 2008). One explanation of this effect is that the interleaved distraction during complex span displaces the memory items from focal attention and requires retrieving them from LTM; in contrast, simple span tasks allow participants to maintain all the items within, and report them directly from, the focus of attention.

Other evidence suggests that the extent to which items are displaced from focal attention depends on the type and amount of distraction. For example, Rose, Buchsbaum, and Craik (2014, 2015) found behavioral and neural evidence for the involvement of LTM in retrieving a single word following just a few seconds of distraction, but the involvement was greater for distraction from a hard- vs. easy-math task. On each trial, a word was encoded with either deep or shallow processing, and then the delay period consisted of either no distraction (rehearsal) or an easy- or hard-series of math problems. For initial recall on the WM task, there was a large levels-of-processing effect following hard math, a smaller effect following easy math, and no effect following rehearsal. For final free recall on the LTM test, the items maintained in the conditions with math were better recalled than the items maintained in the condition without distraction. These results provide evidence for the involvement of LTM in WM tasks following hard math, but not following continuous rehearsal (Rose, Buchsbaum & Craik, 2014; see related behavioral research see Rose & Craik, 2012 and Loaiza & Camos, 2016). These findings were replicated and extended in an fMRI study that supported the interpretations based on differential involvement of frontotemporal networks involved in rehearsal in WM vs. retrieval from LTM (Rose, Craik, & Buchsbaum, 2015; for related research on a hippocampal amnesic patient, see Rose, Olsen, Craik & Rosenbaum, 2012). Thus, these findings collectively suggest that recovering latent items during WM tasks involves retrieval from LTM. In other words, when items are not in the focus of attention, they are in LTM rather than WM (McElree, 2006).

Overall, the literature is clearly mixed regarding whether latent items no longer in focal attention are in an activity-silent state in WM or are simply represented in LTM. The recently developed Dynamic Processing Model of WM (Rose, 2020) attempts to accommodate for this variability. The goal of the present study is to test hypotheses of this model to distinguish between the two accounts regarding whether recovering latent items back into focal attention in WM requires LTM.

The influence of age on reactivating latent WM information

One approach to investigating whether LTM is required to reactivate latent WM items has been to examine differences between groups as a function of development. Compared to young adults, older adults often have deficits in their ability to utilize their attention to maintain information in WM (Bopp & Verhaeghen, 2005; Johnson, Logie, & Brockmole, 2010). One way this age difference in attention has been examined is to assess whether older adults' WM performance benefits less from attentional cues that are presented prior to the onset of the memoranda (i.e., precues) and/or during the delay/maintenance period (i.e., retrocues, as explained previously). There is increasing evidence showing that young adults outperform older adults on WM tasks with precues (Gazzaley, Clapp, Kelley, McEvoy, Knight & Esposito, 2008; Jost, Bryck, Vogel & Mayr, 2011; McNab, Zeidman, Rutledge, Smittenaar, Brown, Adams & Dolan, 2015; but see Souza, 2016). However, some retrocue studies suggest that older adults are able to use retrocues as effectively as young adults to guide their attention to the cued items (Gilchrist, Duarte, & Verhaeghen, 2016; Souza, 2016; Loaiza & Souza, 2018, 2019), whereas others indicate that there are age-related deficits in the use of retrocues for WM tasks (Duarte, Hearons, Jiang, Delvin, Newsome, & Verhaeghen, 2013; Newsome, Duarte, Pun, Smith, Ferber, & Barense, 2015). Thus, whether older adults have a preserved ability to benefit from retrocues is unclear.

The present study aimed to clarify the effects of retrocues on healthy young and older adults' WM and LTM performance in order to examine the ways in which retaining information in different states of prioritization in WM might involve episodic LTM processes and affect subsequent LTM. A related approach that has been used to investigate the role of LTM in WM is to examine the nature of age differences in, and relations among, WM tasks (Hale, Rose, Myerson, Strube, Sommers, Tye-Murray, & Spehar, 2011; Johnson et al.,2010), as well as LTM tests, including subsequent LTM of items initially encoded and maintained in a WM task (Loaiza & McCabe, 2013; Rose, Myerson, Roediger, & Hale, 2010; Rose & Craik, 2012; Rose et al., 2014). This approach has been useful for both WM and cognitive aging researchers because age deficits in WM have been shown to underlie declines in other domains of cognition (Park, 2000), including LTM (McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). The effects of normal aging on LTM have been extensively characterized; for example, age deficits in recognition memory tend to be larger for associative memory (i.e., recognizing the original pairing between two studied items) than for old/new or item recognition memory (i.e., recognizing the individual items irrespective of their association; Old & Naveh-Benjamin, 2008). Therefore, relating older adults' LTM deficits on item and associative LTM tests to WM performance can help reveal the nature of LTM involvement in WM. That is, if reactivating latent WM items requires retrieval from LTM, then healthy older adults who typically exhibit LTM deficits may experience less of a long-term benefit of reactivating these items compared to younger adults.

The current study

In the current study, a double-retrocue WM task that involved recognition of a combination of faces, scenes, or names was administered to examine healthy young and older adults' ability to switch between items in WM (see Figure 1). On each trial, participants viewed two memory items (e.g., a face and place), displayed on the left and right-hand side of the screen, followed by two retrocues interleaved by two test probes. On "stay" trials, the two retrocues pointed to the same memory item (i.e., probes 1 and 2 tested the same memory item), so participants were tested on the item that was to be continuously attended throughout the trial. In contrast, on "switch" trials, the second retrocue pointed to the other memory item (i.e., probes 1 and 2 tested different memory items), so participants had to switch their attention back to the initially uncued item to respond to the second probe. Following all of the trials of the double-retrocue WM task, a subsequent LTM test was administered wherein participants made old/new item-recognition judgments followed by judgments about both the associated location (item-location context memory) and the associated item (item-item associative memory) for each of the items judged ‘old’.

Thus, the current study attempted to replicate and extend aspects of the paradigm and design of LaRoque et al. (2015) Experiments 2 and 3. The current study extends that research in several novel ways. First, we examined whether the effects are similar for healthy young adults as well as a population with LTM deficiencies (i.e., older adults); second, we examined whether the effects were similar for a condition with distraction during the delay and cueing periods; third, we told participants that their subsequent LTM would be tested for each item, its location, and its paired associate (i.e., intentional encoding); and fourth, we examined whether the effects were similar for subsequent LTM tests that assessed both item and associative recognition memory. That is, we assessed the extent to which maintaining items in different states of accessibility may affect not just overall subsequent LTM, but item versus associative memory in particular. Specifically, we measured item-location context memory and item-item associative memory, as well as participants’ confidence in these memory decisions. This allowed us to assess whether the effects of dropping and reactivating an item from focal attention during WM maintenance affects subsequent recollection of the item (i.e., remembering specific details of the associations; Loaiza et al., 2015) in addition to item recognition which is thought to be more heavily influenced by the strength of a familiarity signal (Yonelinas, 2001).

Our two key pre-registered hypotheses centered on the following logic. If recovering latent items involves retrieving them from LTM, then:

1) older adults, who have deficiencies in LTM, should show deficits relative to young adults on the double-retrocue WM task, particularly when trying to reactivate a latent WM item (i.e., probe 2 switch trials; pre-registered hypothesis 1), and

2) both young and older adults’ subsequent LTM should differ between items that were initially held in different states of prioritization on the WM task: A-A, A-U, U-A and U-U condition, particularly between unattended items that were reactivated (U-A condition) vs. items that

were initially attended but dropped from focal attention later on (A-U condition) during the switch trials. That is, if retrieval from LTM is required to reactivate previously uncued items that were not in focal attention, then subsequent LTM for items from the U-A condition should be better than the A-U condition (pre-registered hypothesis 2A). Moreover, if LTM is involved in reactivating U-A items more than maintaining A-U items, then the age difference in subsequent LTM performance should be larger for U-A items than A-U items (pre-registered hypothesis 2B).

To preview the results, in our first experiment we found that WM performance was worse for items that were maintained and reactivated outside of focal attention (probe 2 switch trials) than inside the focus of attention throughout a trial (probe 2 stay trials), and worse for older adults than young adults. That is, we replicated and extended the results of Larocque et al. (2015) in a population with LTM deficiencies (i.e., older adults). To elucidate these results, we designed a second experiment with a classic manipulation to assess the source of differences in memory between the conditions and age groups. In Experiment 2, young adults performed the same task, but while also performing a secondary distractor task during the maintenance and cueing parts of the task trials (i.e., divided attention, DA, to be contrasted with Experiment 1 in which the young and older adults had full attention, FA). For Experiment 2, participants' attention was not divided during the encoding and retrieval phases of the trials in order to ensure that any effects of DA on performance were not due to interference with encoding or retrieval processes. This was to see if dividing young adults' attention during the delay periods would cause them to perform similar to older adults. We hypothesized that control processes are required to focus attention on a cued item to prioritize it over an uncued item; therefore, young adults whose attention is preoccupied by having to perform a secondary distractor task during the cueing and delay period phases should exhibit a similar pattern of errors as the older adult group (cf. Castel & Craik, 2003). Experiment 2 tested and provided support for this hypothesis. However, what is most interesting is how none of the factors that had a strong influence on WM performance (retrocueing, aging, dividing attention) had any reliable or systematic effect on subsequent LTM. As discussed further on, we interpret this to mean that retaining and reactivating latent (unattended) items on the WM task did not involve LTM.

(A) Double-Retrocue Working Memory Task Procedure

(B)

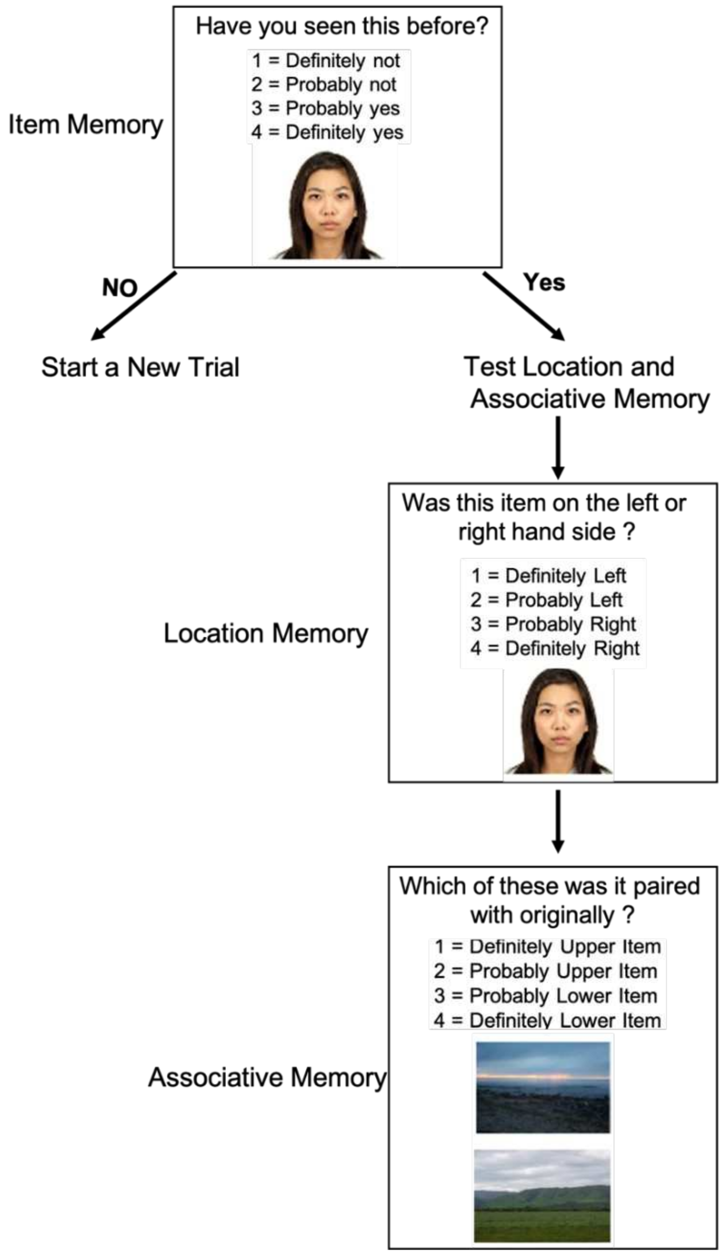


Figure 1. (A) example of the WM task procedure for either a Probe 2 Stay trial (left) or Probe 2 Switch trial (right). For this example of the WM task, if it were a Probe 2 Switch trial, the face (item "A") would be unattended 1st and attended 2nd (UA, and probed by a new item "D") and, thus, the scene (item "B") would be attended 1st and unattended 2nd (AU); if it were a Probe 2 Stay trial, the scene would be attended 1st and attended 2nd (AA, and probed by the old item "B") and, thus, the face would be unattended 1st and unattended 2nd (UU). (B) example of the subsequent LTM recognition test procedure. For this example, if the participant responded "Definitely yes" or "Probably yes" to this "old" item on the Item Memory test, then they were asked to indicate which side of the screen the item was initially presented on, and which item it was initially paired with to test Location and Associative Memory, respectively.

General Methods

Participants

A priori power analyses were conducted using G* Power to estimate the number of participants that would be required in order to detect a reliable effect at least as large as those reported in the prior literature with 95% power and an alpha level of .05. We used the effect sizes reported by Newsome et al. (2015) to show that at least 54 total participants are required to obtain at least 95% power for detecting a reliable age by retrocue effect interaction (assuming a correlation between conditions of .5 and nonsphericity correction at 1). However, because half of the previous studies showed no interaction between age and retrocueing (i.e., Gilchrist et al., 2016; Loaiza & Souza, 2018, 2019; Yi & Friendman, 2014), we aimed to have larger datasets (30 participants for each group) in order to ensure that any failure to detect an effect is not due to undersampling or having insufficient power for this study.

For Experiment 1, 30 young [23 females, aged 18 to 21 yrs (M = 19.87, SD = 1.19)] and 30 older adults [16 females, aged 64 to 81 yrs (M = 71.92, SD = 3.97)] participated. For Experiment 2, a total of 30 young adults[23 females, aged 18 to 24 yrs (M = 20.10, SD = 1.37)] participated. All participants had normal or corrected-to-normal vision and hearing, the ability to discriminate between the colors red and green, and have used English as their primary language for at least 15 years.

All participants were screened for the presence of possible neurocognitive dysfunction with the Telephone Interview of Cognitive Status (TICS, Knopman, Roberts, Geda, Pankratz, Christianson, Petersen, & Rocca, 2010). All participants had a modified-TICS score greater than 34 suggesting that all participants had normal neurocognitive function (Knopman et al., 2010). The mean scores on the modified-TICS for the young adults in Experiments 1 and 2 were 40.47 (SD = 3.22) and 40.48 (SD = 2.95), respectively, and the mean score for the older adults was 38.15 (SD = 2.94). Performance was significantly higher for the young adults in both Experiments 1 and 2 compared to the older adults ($t(55) = 2.82$, $p = 0.007$ and $t(54) = 2.96$, $p = 0.004$, respectively). Specifically, young adults in both experiments outperformed older adults on the initial free recall test ($t(55) = 3.25$, $p = 0.001$ and $t(54) = 3.92$, $p = 0.0003$, respectively), and also on the final free recall test ($t(55) = 3.15$, $p = 0.002$ and $t(54) = 4.21$, $p = 0.0001$, respectively) on the modified-TICS. These results confirmed that the older adult group had deficits in episodic LTM compared to the young adults.

Participants were compensated with either extra course credit or a gift card ($15 per hour) for their participation. This protocol was approved by the University of Notre Dame's Institutional Review Board (Protocol # 18-01-4374).

Materials and Procedure

The experimental tasks were programmed in PsychoPy2 (Pierce, Gray, Simpson, MacAskill, Höchenberger, Sogo, Kastman, & Lindeløv, 2019) for in-person or online administration via Pavlovia.org. The task code is available at (https://osf.io/cnvma/). Participants completed the experiment either in-person in our laboratory while abiding by the University- and IRB-approved COVID-19 safety protocols (10 Young, 12 Old for Exp. 1; 20 Young for Exp. 2) or online with an experimenter delivering the instructions and practice trials, and supervising the completion of the experimental trials via Zoom session with video-recorded screen sharing (20 Young, 18 Old for Exp. 1; 10 Young for Exp. 2). Control analyses were conducted to see if the location of testing interacted with the factors of interest. For each group, there were no interactions between probe type and location of testing (see Supplemental Material).

During the instructions phase, all participants were told that they would perform a "short-term memory" task followed by a "long-term memory" test of their memory for the items from the initial short-term memory test. That is, in contrast to LaRocque et al. 2015, this study involved intentional, not incidental, encoding for the subsequent LTM test. Participants were specifically told that their memory for each item from the WM task would be tested later on in the session, including their memory for the item's location (whether it was presented on the left or right side of the screen), and the other item with which it was paired.

A diagram of an example trial of the double-retrocue WM task with two retrocues and two recognition probes on each trial is shown in Figure 1. In total, there were 48 trials. Participants were told to look at the central fixation cross at the start of each trial. Then two items from different categories were presented -- a combination of pictures of either a face, a name, or a scene, with one to the left of central fixation (item A in the task diagram in Figure 1), and the other to the right of central fixation (item B).

The pair of stimuli on each trial was presented for 2 seconds, then there was an arrow presented in the center of the screen that pointed to either the left or right item to cue participants that this item would be tested first. After a short delay, participants saw a probe stimulus (item C) that was either an exact match (the exact same image as the cued stimulus) or a mismatch (an image of a novel item from the same category as the cued item). Participants had 2 seconds to determine whether the cued item (A or B) and the probe (C) were the same or different. If they were the same, participants pressed 1 on the computer's T9 keypad with their right pointer finger as quickly as possible; if they were different, participants pressed 2 with their right middle finger as quickly as possible. If their responses were correct, the fixation cross turned green; otherwise, the fixation cross turned red. Then participants saw a second arrow that pointed to either the left or right item to cue participants to the item (the initially presented A or B item) that would be tested second. After another short delay, participants saw a second probe stimulus (item D) that was either a match (the same as the initially presented cued stimulus) or a mismatch (different from the initially presented cued item). Participants needed to decide if the probe was a match or nonmatch of the initially presented stimulus at that location (not a match of the first test probe).

For Experiment 2, participants performed the same task, but with a secondary odd-even digit-parity task during the cue and delay periods of the WM task to divide their attention from maintaining the items in focal attention throughout the cue and delay periods. Immediately after the two stimuli were presented, participants were to attend to a random series of digits (1 to 9) presented auditorily through headphones at a comfortable listening level (~50% of PC volume); participants were to press the 'o' key with their left middle finger for odd digits and the 'p' key with their left pointer finger for even digits as quickly and as accurately as possible. Participants received feedback immediately following each response by the presentation of either a high (800 Hz) tone for a correct response or a low (400 Hz) tone for an incorrect response. The rate at which the digits were presented was individually determined during a pretesting procedure that titrated the level of difficulty across participants. Before the participant performed the test trials for the WM task with the secondary odd-even digit-parity task, they performed the odd-even digit task on its own to titrate the rate of presentation. The participant was to press the correct response key as quickly as possible to indicate whether the digit was odd or even; the rate of presentation was adjusted based on the accuracy of the response according to a staircase method using a one-up and three-down rule. The initial responses were to be entered within 1 second from the onset of the auditorily presented digit. If there were three successive correct responses, the response window for the next trial was decreased by 20%. If an incorrect response was made, the response time window for the next trial was increased by 25% (i.e., the current time window * 1.25 in seconds). The up/down staircase procedure ended after 10 reversals, and the titrated presentation rate for the participant was applied to the WM task code for the test trials by calculating the average response time of the last three reversals from the pretest and adding two standard deviations. We first confirmed that young adults in the DA group performed the DA task at a high level before proceeding with the WM test trials. Average accuracy on the secondary task during the WM test trials was 81.03% correct (SD = 6.4%). Based on participants' performance on the odd-even digit task during the pretesting procedure, it was possible for different numbers of digits to be presented during the cue and delay periods of the WM task test trials. For example, as shown in Figure 1, the time between the target and probe was 4 sec. If the titrated presentation rate was above 1.2 sec, then there were two digits presented during the delay period, otherwise there were three digits. There were 28 participants with two digits presented during each cue and delay period and 2 participants with three digits. Control analyses were conducted to confirm that the pattern of results did not differ if the participants with three digits were excluded (see Supplemental Material). Pilot testing of the procedure suggested that this formula would result in participants' performance on both the odd-even digit task and the WM task being off of both ceiling- and floor-level performance for participants who were equally prioritizing both tasks. All other details were the same as Experiment 1.

Following the WM task for both Experiments, participants were administered the TICS, which served as both a screening measure of neuropsychological function and a distractor task performed after the WM and before the LTM test to ensure that no items from the WM test were continuously maintained until the LTM test. The TICS lasted approximately 5 minutes. Then participants were told that they would take the subsequent LTM test in which half of the ‘old’ items from the initial 48 trials of the WM task, and an equal number of novel, lure items from each category (48 in total) would be presented one at a time in a random order. For those items from the initial WM task, there were an equal number of ‘old’ items from the four different conditions of the WM test trials (i.e., AA: the tested item was attended 1st and attended 2nd; AU: the tested item was attended 1st and unattended 2nd; UA: the tested item was unattended 1st and attended 2nd; UU: the tested item was unattended 1st and unattended 2nd). The old items were also equally balanced for each stimulus category and for items that had appeared on the left or right side of the screen.[2]

For each image, the participant was asked to indicate whether they thought the item was ‘old’ (a to-be-remembered item from the WM test) or ‘new’ (not presented on the WM test) on a four-point confidence scale (definitely old, probably old, probably new, or definitely new). For this measure of item memory, the proportion of old items called definitely or probably old was recorded as hits, the proportion of old items called definitely or probably new was recorded as misses, the proportion of new items called definitely or probably old was recorded as false alarms, and the proportion of new items called definitely or probably new was recorded as correct rejections for each condition. Item memory accuracy was scored as hits - false alarms. When participants judged an item to be ‘old’, they were then asked to first indicate the side of the screen on which the item was initially presented using the same kind of four-point confidence scale (definitely left, probably left, probably right, or definitely right) to measure their location memory; then they were asked to indicate which one of two images (one presented in the upper- and one in the lower-half of the screen) was the corresponding item that was originally presented and maintained with the item in question during the WM task; one image was the item that was initially presented with the item in question and the other was a novel, lure item from the same category that was never presented in the experiment; they were to indicate their response with the same kind of four-point confidence scale (definitely the upper item, probably the upper item, probably the lower item, or definitely the lower item).

2 Note that LaRocque et al. (2015) included several control conditions to rule out potential confounding effects of attentional cueing and testing on subsequent LTM. Because their two experiments already ruled out these confounds, we did not include such trials in the current experiments.

For the stimuli (pictures of faces, scenes, and names), face stimuli were obtained from the Chicago Face Database (Singh, Gambrell & Correll, 2022). Faces with neutral expressions were selected to balance gender (male and female) and race (White, Black, Latin, Asian). The scene stimuli were obtained from the Place365 dataset (Zhou, Lapedriza, Khosla, Oliva & Torralba, 2017). Scenes that would not be readily identifiable or recognized by our participants were selected. The name stimuli were selected from the US First Names Database (https://data.world/len/us-first-names-database) which is a publicly available database of commonly used names in the United States. Thus, the names were all relatively familiar, short, and easy to pronounce/rehearse for our American dwelling, English speaking participants. The full set of stimuli used are available at https://osf.io/tmxfc/.

Data Analysis and Predictions

With regard to the WM task, accuracy values on probe 1, probe 2 stay and probe 2 switch trials were the primary dependent measures of interest. Response time (RT) data and analyses are reported in the supplemental materials for the interested reader. Note that because stay and switch trials were equally probable (50%), it was unnecessary to distinguish between stay and switch trials for probe 1 responses. We predicted that accuracy would be better for both probe 1 and probe 2 stay trials than probe 2 switch trials because these probes tested the item that was continuously retained in focal attention throughout the trial whereas probe 2 switch trials probed memory for the passively-retained, "latent" item held outside of focal attention. Note that, despite the longer retention interval on probe 2 stay vs. probe 1, performance on probe 2 stay trials was not expected to be worse than probe 1 because the same item was tested on probe 2 and feedback was provided after the probe 1 response; however, there is one exception: on probe 2 stay trials in which the *same nonmatch* probe was presented for both probe 1 and probe 2, we expected that these repeated lures would be particularly difficult to reject as a nonmatch of the target item because the (nonmatch) probe 2 matched the (nonmatch) probe 1 . That is, the same nonmatch probe stimulus was presented on both probe 1 and probe 2 for this subset of stay trials, which results in a higher degree of interference and greater difficulty with rejecting these repeated negative probes (Jonides & Nee, 2006). These probes were similar to 'recent negative probes' on the n-back task, which have been shown to be exceptionally hard to reject due to their increased interference, especially for older adults (Jennings & Jacoby, 2003). Investigation of these particular trial types is the focus of a separate study to test the a priori prediction that performance should be worse, and age differences larger, for these trials relative to probe 2 stay trials with non-repeated lures (see the pre-registered hypotheses). For present purposes, these trial types were not included in the analyses reported below.

With regards to age differences in WM, given the extensive literature on the topic, we predicted that WM accuracy would be better on average for the young adult group than the older group overall (main effect of age on accuracy). We also expected that there would be an interaction between age and trial type due to age differences in the use of retrocues to prioritize information held in WM. If reactivating latent items involves retrieval from LTM, then older adults, who have LTM deficiencies, should perform worse than young adults with FA on probe 2 switch trials overall. Because older adults should be able to maintain 1 or 2 items in focal attention as well as young adults with FA, there should be a small or nonexistent age difference on probe 1; that is, the age difference on probe 1 and on probe 2 stay trials should be smaller than the age difference on probe 2 switch trials (see the pre-registered hypotheses).

With respect to the LTM test, we predicted that memory performance would be better for young than older adults overall, particularly for location memory (whether the item was initially presented on the left or right side of the screen) and associative memory (which items were paired together on a trial). Because age differences are often small or nonexistent for item recognition memory (Old & Naveh-Benjamin, 2008), we predicted that age differences should be smaller for item memory (old-new recognition decisions about which items had been seen before in the session) than location or associative memory.

To test these hypotheses, subsequent long-term recognition memory was compared for items initially held in the four conditions (AA, AU, UA, UU) for the two age groups using a mixed repeated-measures analysis of variance (ANOVA). Follow up t-tests were used to test the a priori hypotheses described previously. These analyses were conducted separately for WM accuracy for probe 1, probe 2 stay, and probe 2 switch trials, and for old/new item recognition, item-location context memory, item-item associative memory (collapsed over ‘definitely’ and ‘probably’ judgments). The analyses on LTM were also repeated to examine performance separately for items recognized with high (‘definitely’) vs. low (‘probably’) confidence (see supplemental results).

Note that the most appropriate contrast for subsequent LTM was between UA items and AU items because UA items were the items that were potentially dropped from focal attention and retrieved with LTM processes whereas AU items were continuously retained and reported directly from focal attention. Comparisons to AA and UU items are complicated by the fact that AA items were tested twice, potentially receiving twice the benefit of retrieval practice as AU or UA items. The UU items were not tested at all during the WM test; therefore, they cannot provide insight about the nature of WM retrieval and its impacts on subsequent LTM, but they can provide a baseline level of LTM performance for items that were never tested to compare to LTM performance for UA and AU items. If LTM is better for UA than AU items it would suggest that reactivating UA items involved more episodic LTM retrieval practice than AU items. However, if LTM does not differ between UA and UU items, it would suggest that the UA items did not get any benefit of episodic LTM retrieval practice. Also note that, as in previous research (e.g., Rose & Craik, 2012; Rose et al., 2014), we analyzed those items that were initially retrieved correctly on the WM test in order to adequately compare the retention rates for items from the different WM conditions and between younger and older adults’ LTM, whose WM accuracy rates were expected to differ. Otherwise, it would remain unclear whether differences in LTM performance were due to differences in initial WM accuracy or the ways in which WM maintenance or aging affects subsequent LTM.

Data exclusion criterion

Video of the experiment sessions were recorded to monitor the participants’ level of arousal, eye blinks and movements during the stimulus, cue, or probe presentation periods of the task, and also to see if any interruptions or excessively long pauses impacted data collection. No data needed to be excluded on the basis of these criteria. The recorded experimental sessions were also examined to see if the participant did not understand or follow the instructions (e.g., they reversed the mapping of the response buttons) and, therefore, should be excluded from analyses. Three participants from the older adults group in Experiment

1 were excluded because their average WM accuracy in a condition was less than 55% and, upon review of the recorded session, it was apparent that they did not understand or follow the instructions. For Experiment 2, if a participant's average accuracy was below either 55% in a condition of the WM task or 70% on the secondary odd-even digit task, the data were excluded from the analyses. One participant was excluded due to low performance on the secondary task.

All data have been made publicly available and can be accessed at CurateND.edu. This study's design and hypotheses were pre-registered; see https://osf.io/z9cgq/

Results

To assess the role of LTM in WM, the subsequent LTM data of items processed on the WM task are of primary interest. However, we first report the WM performance data to contextualize interpretation of the LTM data.

WM Results

Experiment 1. Average performance on the WM task for the young and older adult groups with FA for the three types of trial probes are shown in Figure 2, along with the average performance for the young adult group with DA in Experiment 2 for comparison.

To test our pre-registered hypothesis 1, we compared age differences on the items that were possibly recovered via LTM retrieval processes (probe 2 switch trials, UA items). Age differences on these items were compared to age differences on the items that were presumably maintained in focal attention throughout the trial (probe 2 stay trials, AA items) because these. If recovering items on probe 2 switch trials involved retrieval from LTM, which is deficient in older age, then the age difference should be larger on probe 2 switch trials compared to probe 2 stay trials that presumably do not require retrieval from LTM. This analysis was conducted with a repeated-measures ANOVA on accuracy with group (Young FA vs. Old FA) as a between-subjects factor and probe-type (probe 1, probe 2 stay, probe 2 switch) as a within-subjects factor. Against this hypothesis, however, the size of the age difference on the probe 2 switch trials (Young FA mean - Old FA mean = 0.06; 95% CI: 0.018 to 0.109) was similar to that of probe 2 stay trials (Young FA mean - Old FA mean = 0.06; 95% CI: 0.03 to 0.09).

The ANOVA did show significant main effects of age group [$F(1,55) = 10.06$, $p = 0.002$, $\eta p2 = 0.16$, $BF_{10} = 15.36$] and probe-type [$F(2,110) = 16.13$, $p < 0.001$, $\eta p2 = 0.23$, $BF_{10} = 8347.06$]. The main effect of age group was significant because, as expected, WM performance was better for young with FA than older adults overall. There was a significant interaction between these two factors [$F(2,110) = 5.03$, $p = 0.008$, $\eta p2 = 0.08$, $BF_{10} = 7.67$]. Simple main effect analysis showed that accuracy was higher on probe 2 stay trials than probe 2 switch trials for both young [$t(29) = 4.65$, $p < 0.001$, $BF_{10} = 368.55$] and older adults [$t(26) = 3.29$, $p = 0.003$, $BF_{10} = 13.57$], suggesting that young and older adults with FA were able to use the retrocues to prioritize the cued item over the uncued item. However, the pattern suggests that older adults used the retrocues differently from young adults with FA. WM accuracy for young adults with FA benefitted between probe 1 and probe 2 on stay trials, likely due to being tested on the same item twice and receiving feedback following probe 1 [$t(29) = 6.55$, $p < 0.001$, $BF_{10} = 45499.43$]; in contrast, older adults did not benefit from repeated testing with feedback [$t(26) = 0.56$, $p = 0.58$, $BF_{01}=4.26$]. Instead, the retrocue effect for the older adults occurred because they performed worse on probe 2 switch trials compared to probe 1 [$t(26) = 3.57$, $p =0.001$, $BF_{10} = 25.20$], whereas there was no difference between probe 1 and probe 2 switch accuracy for young adults with FA [$t(29) =0.66$, $p =0.51$, $BF_{01}=4.20$]. There was no effect of age group (young FA vs. old FA) on probe 1 [$t(55) = 1.12$, $p = 0.268$, $BF_{01}= 2.21$], but there were age-related deficits on probe 2 stay [$t(55) = 4.03$, $p < 0.001$, $BF_{10} =136.91$] and probe 2 switch [$t(55) = 2.86$, $p = 0.006$, $BF_{10} = 7.18$] trials. However, as indicated previously, the age difference was similar between probe 2 stay and switch trials, which is inconsistent with an LTM account of latent WM retrieval.

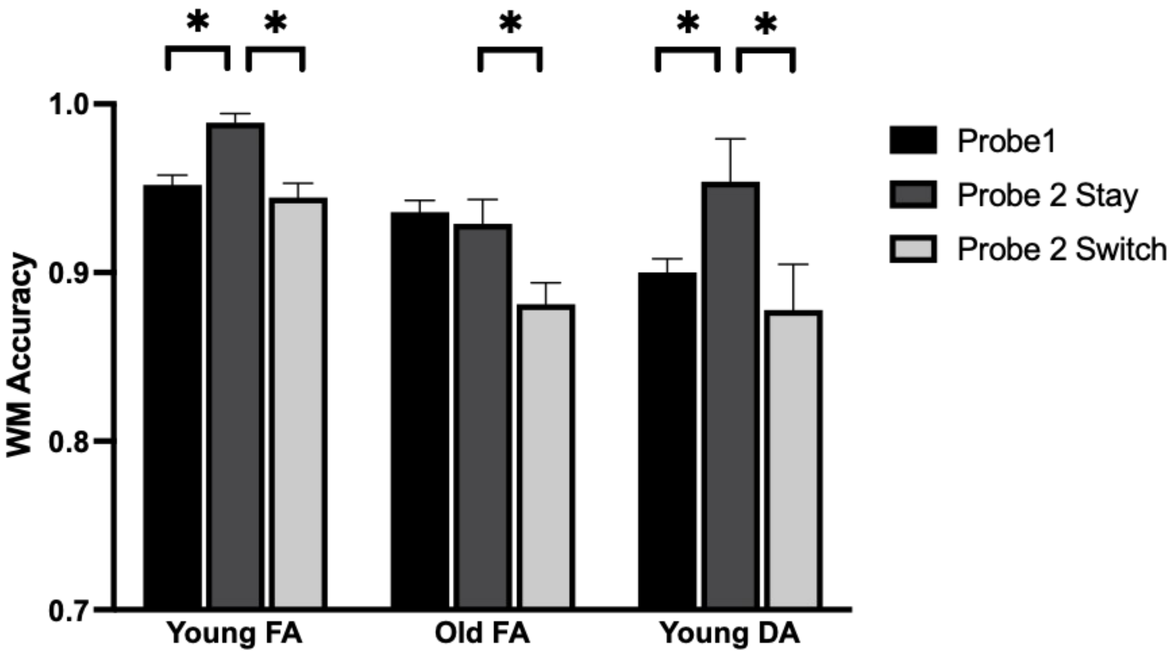


Figure 2. Average accuracy on the WM task for the three types of trial probes in Experiment 1 (Young FA and Old FA) and Experiment 2 (Young DA). FA = full attention; DA = divided attention. Error bars indicate the standard error of mean. *indicates *p* < .05 from two-tailed, paired-sample t-tests.

Experiment 2. Average performance on the WM task for young adults with DA during the cue and delay periods for the three types of trial probes are shown on the right side of Figure 2 alongside the data for the Young FA and Old FA groups from Experiment 1 for comparison. A repeated-measures ANOVA on accuracy with age-group (Young DA vs. Old FA) as a between-subjects factor and probe-type (probe 1, probe 2 stay, probe 2 switch) as a within-subjects factor showed no main effect of age group [$F(1,54) = 0.08$, $p = 0.78$, $\eta p2 = 0.001$, $BF_{01} = 3.56$], suggesting that performance was not better overall for the young DA group than the older adult FA group. There was a significant main effect of probe-type [$F(2,108) = 17.77$, $p < 0.001$, $\eta p2 = 0.25$, $BF_{10} = 41546.37$], and the interaction between age group and probe-type was significant [$F(2,108) = 4.20$, $p = 0.017$, $\eta p2 = 0.07$, $BF_{10}= 2.89$]. The group difference (old FA vs. young DA) was significant on probe 1 trials [$t(54) = 2.26$, $p = 0.028$, $BF_{10} = 2.141$], but was not on either probe 2 stay trials [$t(54) = 1.49$, $p =0.142$, $BF_{01} = 1.477$] or probe 2 switch trials [$t(54) = 0.12$, $p = 0.905$, $BF_{01} = 3.68$]. The size of the group difference on probe 2 switch trials (Old FA mean - Young DA mean = 0.003; 95% CI: -0.05 to 0.06) was similar to that of probe 1 trials (Old

FA mean - Young DA mean = 0.035; 95% CI: -0.004 to 0.07) and probe 2 stay trials (Old FA mean - Young DA mean = -0.02; 95% CI: -0.06 to 0.008).

For completeness, exploratory analyses were carried out to compare WM performance between the young DA group and young FA group. The young DA group had higher accuracy on probe 2 stay trials than both probe 1 [$t(28) = 4.85$, $p < 0.001$, $BF_{10} = 573.49$] and probe 2 switch trials [$t(28) = 4.37$, $p < 0.001$, $BF_{10} = 175.57$]. The ANOVA comparing young DA to young FA showed a main effect of probe-type [$F(2,114) = 24.91$, $p < 0.001$, $\eta p2 = 0.304$, $BF_{10} = 8096429$] anda main effect of group [$F(1,57) = 15.24$, $p < 0.001$, $BF_{10} = 93.43$], but there was no interaction between the two factors [$F(2,114) = 1.59$, $p = 0.208$, $\eta p2 = 0.027$, $BF_{01} = 2.69$]. These results suggest that, although young adults with DA performed more poorly overall than young adults with FA, they were still able to similarly use both the retrocues to prioritize the cued items and the feedback following probe 1 to benefit their performance on probe 2 stay trials.

Analysis of young and older adults' mean response times (RTs) for the different trial types generally converged with interpretation of the accuracy data. For the sake of brevity, interested readers are referred to the supplemental materials for the RT data and analyses.

Subsequent LTM Results

Average performance on the subsequent item, location, and associative LTM recognition tests for correct items from the initial WM test in the different conditions (AA, AU, UA, UU) are shown in Figure 3 for the young and older adults with FA and young adults with DA.

Experiment 1. To test our pre-registered hypothesis 2A, we compared the mean difference between items that were dropped from focal attention and recovered (UA) and items that were retained in focal attention and then dropped from maintenance (AU) on the hit - false alarm rate for the item, location, and associative recognition tests. The mean overall differences (UA - AU), for item, location, and associative recognition were -0.001 (95% CI: -0.07 to 0.07), 0.03 (95% CI: -0.09 to 0.03), and -0.04 (95% CI: -0.02 to 0.10), respectively. Thus, subsequent LTM was not better for UA than AU items overall.

To test our pre-registered hypothesis 2B, we compared the mean age difference between the young and older adult groups on both UA and AU items to test the hypothesis that LTM was involved more in the former than the latter condition. The mean age difference was similar between UA and AU items on the item memory test (UA: Old FA mean - Young FA mean = -0.03; 95% CI: -0.08 to 0.13; AU: Old FA mean - Young FA mean = -0.03; 95% CI: -0.12 to 0.06), location memory test (UA: Old FA mean - Young FA mean =-0.04; 95% CI: -0.05 to 0.13; AU: Old FA mean - Young FA mean = 0.001; 95% CI: -0.09 to 0.09), and associative memory test (UA: Old FA mean - Young FA mean = -0.04; 95% CI: -0.05 to 0.13; AU: Old FA mean - Young FA mean = -0.02; 95% CI: -0.07 to 0.11).

These contrasts were established from the results of mixed, repeated-measures ANOVAs for each type of memory test with condition (AA, AU, UA, UU) as a within-subjects factor and age group (Young FA vs. Old FA) as a between-subjects factor. With regard to the overall main effects and interactions from these ANOVAs, as may be seen in Figure 3, for item memory (top), there were no significant main effects of either condition [$F(3,165) = 2.57$, $p = 0.056$, $\eta p2 = 0.045$, $BF_{01}= 2.82$] or age group [$F(1,55) = 0.095$, $p = 0.759$, $\eta p2 = 0.002$, $BF_{01}=3.21$]. The interaction between these two factors also was not significant [$F(3,165) = 1.31$, $p = 0.266$, $\eta p2 = 0.023$, $BF_{01}= 4.86$]. These results suggest that the way the items were initially maintained and retrieved on the WM task did not influence subsequent item recognition memory, and also that item memory did not differ between the young and older adults. For location memory (Figure 3 middle), the main effects of both condition [$F(3,165) = 0.62$, $p = 0.604$, $\eta p2 = 0.011$, $BF_{01}= 21.71$], and age group [$F(1,55) = 1.19$, $p = 0.28$, $\eta p2 = 0.021$, $BF_{01}= 2.93$] were not significant; neither was the interaction [$F(3,165) = 1.04$, $p = 0.375$, $\eta p2 = 0.019$, $BF_{01}= 6.09$]. For associative memory (Figure 3 bottom), the main effect of age [$F(1,55) = 0.39$, $p = 0.54$, $\eta p2 = 0.007$, $BF_{01} = 4.80$] and the age by condition interaction were not significant [$F(3,165) = 0.66$, $p = 0.59$, $\eta p2 = 0.012$, $BF_{01} = 9.82$]. The main effect of condition was significant [$F(3,165) = 2.70$, $p = 0.048$, $\eta p2 = 0.047$, $BF_{10} = 0.88$] because memory was better for AA items, that were tested twice, compared to UU items, that were never tested.

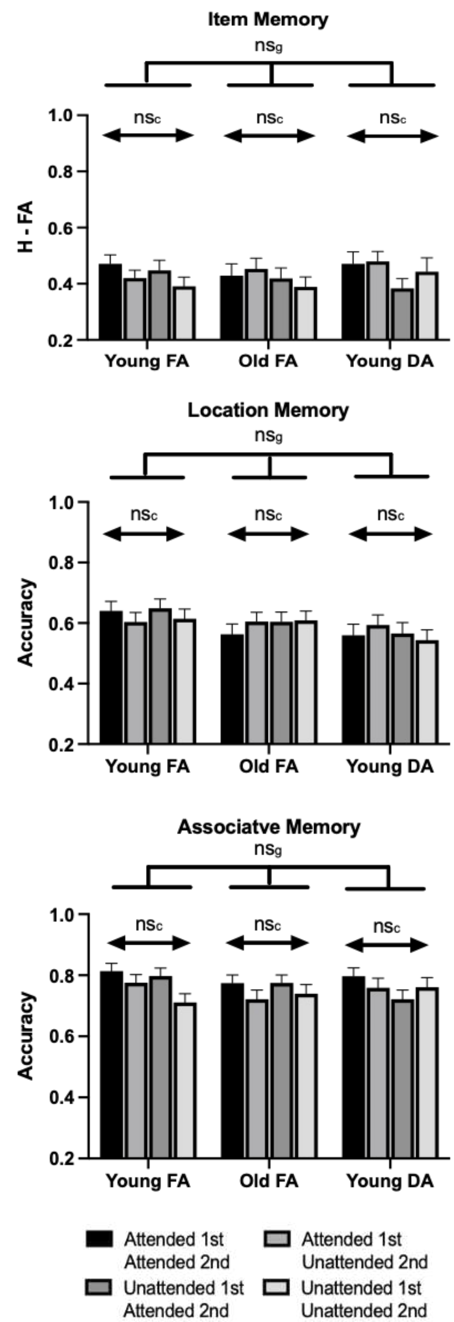


Figure 3. Average performance on the subsequent LTM tests of item (top), location (middle), and associative (bottom) recognition memory for items held in the attended or unattended conditions during the initial double-retrocue WM task for the Young FA, Old FA and Young DA groups (FA =full attention; DA = divided attention). Error bars are the standard error of mean. ns_c indicates a non-significant effect of condition (collapsing over groups); ns_g indicates a non-significant effect of group (collapsing over conditions); the interactions were also not significant (see text).

Experiment 2. Adding the Young adult group with DA during the maintenance and cueing phases of the WM task allowed us to assess the effect of dividing attention on subsequent LTM. The Young DA group was compared to the old FA group first with a series of mixed, repeated-measures ANOVAs for each type of memory test with condition (AA, AU, UA, UU) as a within-subjects factor and group (Old FA vs. Young DA) as a between-subjects factor on performance for the item, location, and associative recognition tests. For item memory, there was no main effect of group [$F(1,54) = 0.001$, $p = 0.97$, $\eta p2 < 0.001$, $BF_{01} = 3.36$], nor an interaction with condition [$F(3,162) = 0.58$, $p = 0.63$, $\eta p2 = 0.011$, $BF_{01} = 10.95$], but there was a main effect of condition [$F(3,162) = 3.39$, $p = 0.019$, $\eta p2 = 0.059$, $BF_{10} = 1.62$]. This main effect was due to item recognition being better for AA, which were tested twice with feedback, than UU items, which were never tested (i.e., a retrieval practice/testing effect). For location memory, there were no main effects of condition [$F(3,162) = 0.78$, $p = 0.51$, $\eta p2 = 0.014$, $BF_{01} = 16.31$] or group [$F(1,54) = 0.72$, $p = 0.40$, $\eta p2 = 0.013$, $BF_{01} = 4.14$], and the interaction was not significant [$F(3,162) = 0.50$, $p = 0.69$, $\eta p2 = 0.009$, $BF_{01} = 11.39$]. For associative memory, there were no main effects of condition [$F(3,162) = 1.15$, $p = 0.33$, $\eta p2 = 0.021$, $BF_{01} = 10.09$] or group [$F(1,54) = 0.03$, $p = 0.87$, $\eta p2 = 0.016$, $BF01 = 4.92$], and the interaction was not significant [$F(3,162) = 1.24$, $p = 0.30$, $\eta p2 = 0.022$, $BF_{01} = 4.91$].

Next, for completeness, exploratory analyses were carried out to compare performance between the young DA group and the young FA group. For item memory, there were no main effects of condition [$F(3,171) = 2.60$, $p = 0.054$, $\eta p2 = 0.044$, $BF_{01} = 2.12$] or group [$F(1,57) = 0.096$, $p = 0.76$, $\eta p2 = 0.002$, $BF_{01} = 3.52$], and the interaction was not significant [$F(3,171) = 2.43$, $p = 0.07$, $\eta p2 = 0.041$, $BF_{01} = 1.41$]. For location memory, there were no main effects of condition [$F(3,171) = 0.28$, $p = 0.84$, $\eta p2 = 0.005$, $BF_{01} = 3.23$] or group [$F(1,57) = 4.00$, $p = 0.050$, $\eta p2 = 0.065$, $BF_{01} = 1.20$], and the interaction was not significant [$F(3,171) = 0.69$, $p = 0.56$, $\eta p2 = 0.012$, $BF_{01} = 9.06$]. For associative memory, there were no main effects of condition [$F(3,171) = 2.10$, $p = 0.10$, $\eta p2 = 0.035$, $BF_{01} = 3.24$] or group [$F(1,57) = 0.56$, $p = 0.46$, $\eta p2 = 0.010$, $BF_{01} = 4.17$], and the interaction was not significant [$F(3,171) = 2.17$, $p = 0.09$, $\eta p2 = 0.037$, $BF_{01} = 1.66$].

In summary, although dividing attention during WM maintenance and cueing affected WM performance, it did not change participants' performance on the item, location, or associative LTM tests. Indeed, LTM performance was similar regardless of the group and the ways in which the items were attended or unattended during the original WM task.

Exploratory analyses on high- and low-confidence judgments. Our pre-registered hypotheses did not specify any a priori predictions about differences in the effects of retrocueing, age, or divided attention on subsequent LTM as a function of confidence. However, we conducted these exploratory analyses on item, location and associative memory separated by high and low confidence judgments, and reported them in the Supplemental Material to help guide future research for the interested reader. An important caveat to interpreting these analyses on the accuracy of responses is that there may be differences in the numbers of high and low confidence responses that the young and older adults provided for items from each condition. To summarize the results, there were no group by condition interactions that differed for high vs. low confidence judgments; the only exception was an interaction on the item memory test because the hit-FA rate for high confidence judgments was higher on UU items, and lower for AA items, for young than older adults, respectively (see Supplemental Material). There were no age differences in mean hit-FA rates between the critical AU or UA items for high and low confidence judgments on the item memory test, and there were no other group by condition interactions for either location or associative LTM that differed for high and low confidence judgements. Therefore, the main conclusions regarding LTM in this paper largely do not need to be qualified by these analyses separated by the level of confidence.

General Discussion

This study was conducted to assess the role of LTM retrieval processes in retaining and reactivating an item outside the focus of attention in WM. To assess this, we compared young and older adults' performance on subsequent LTM for the items that were initially held in different states during the WM task. Three factors affected performance on the WM task: WM was worse for 1) items that were retained and reactivated outside of focal attention (Probe 2 switch trials vs. Probe 2 stay trials), 2) older adults vs. young adults, and 3) young adults with divided vs. full attention. Importantly, WM performance was worse for probe 2 switch than stay trials, suggesting that both younger and older adults retained these items outside of focal attention. This result is important to establish that the items were either held in a latent state, but still in WM (as the activity-silent account predicts) or may have required LTM retrieval processes to reactivate them (as the LTM-retrieval account predicts). If the latter account is correct, that latent items must be reactivated from LTM to return them to focal attention, then a subsequent test of LTM should show greater performance for the reactivated items (i.e., UA items) compared to the items that were initially attended, but then were dropped from WM (i.e., AU items). Although cueing, age, and divided attention all had substantial effects on initial WM performance, none of these factors had any systematic effect on subsequent LTM performance regardless of whether LTM was measured by item memory, location memory, or associative memory, with either high or low confidence.[3] These results replicate and extend those of LaRocque and colleagues (2015): Reactivating latent items in WM does not require retrieval from LTM in either older adults or younger adults under full or divided attention, even when participants intentionally (rather than incidentally) encode the items for an upcoming LTM test that included associative judgments and confidence ratings (rather than two-alternative forced-choice judgments). Thus, several pieces of evidence suggest that retaining and reactivating items on this WM task did not implicate episodic LTM retrieval processes. We discuss this interpretation and its implications for WM theory below.

No impact of switching attention in WM on LTM

On a WM task, if an item is dropped from focal attention and is represented in LTM, then retrieving it would involve retrieval practice and benefit subsequent LTM more than an item that is continuously retained in focal attention (i.e., a McCabe effect; McCabe, 2008). Thus, LTM should be greater for reactivated, latent (UA) items compared to items that were initially attended, but dropped from focal attention (AU) during the switch trials (pre-registered prediction 2A). The results from this double-retrocue procedure showed that subsequent LTM was generally not affected by the way that participants maintained and reactivated the items on the initial WM test; performance was similar regardless of the different states of prioritization on the WM task. This conflicts with the prediction from the LTM-retrieval account

3 Although initial WM was worse for these items, especially for older adults and young adults with DA, we were able to fairly compare subsequent LTM for these items by focusing on those item types that were initially remembered correctly on the WM test.

that latent items are retained in LTM and reactivated with episodic retrieval processes during the WM task.

As described in the introduction, this study was motivated by the findings of Lewis-Peacock et al. (2012) and Rose et al. (2016) which administered a very similar double-retrocue task while recording and decoding patterns of fMRI and EEG activity associated with the categories of items that were cued, uncued, or absent on a given trial. Decoding accuracy detected elevated activation of both items when they were initially presented and held in WM. But, when a retrocue signaled which item was to be tested on the first probe, neural evidence for the uncued item dropped to baseline as if it were no longer being actively retained in focal attention. This occurred despite the fact that this item was technically still "in WM'' and could be rapidly and accurately returned to focal attention as reflected by a return of neural decoding. One interpretation of this phenomenon was that, when the uncued item was dropped from focal attention, it was represented in LTM. This interpretation is consistent with several prominent models of WM (Baddeley, 1986; Cowan, 1999; McCabe, 2008; Unsworth & Engle, 2007): Regardless of how distinct WM is from LTM for each model, all would agree that information no longer active in WM necessarily must be stored in and retrieved from LTM.

Indeed, Foster et al. (in press) recently suggested that a lack of neural activation of an item during the delay period of a WM task indicates that the item is not in WM, but in LTM instead. Using a double-retrocue task identical to those that have shown the return to baseline pattern of decoding of the uncued item (Lewis-Peacock et al., 2012; LaRocque et al., 2013; Rose et al., 2016), we tested whether the uncued item was in LTM and recovered with episodic LTM retrieval processes when subsequently cued as the target item (for probe 2 switch trials). Given that there were only two items to retain on each trial, the embedded process model of WM (Cowan, 1999, 2001) assumes that both items should be retained within the three to four item maximum capacity of the focus of attention throughout the trial and reported with high accuracy. However, items were remembered more poorly on switch trials than on stay trials, which suggests that items were dropped from focal attention and then recovered on switch trials. According to a single item focus model (McElree, 1998), these "latent" items were represented in and recovered via LTM processes so, subsequent LTM should have been better for items recovered on switch trials (UA) than items that were retained in focal attention, dropped and not recovered again on the WM task trial (AU items). We found no evidence to support this hypothesis.

Instead, the current data seem to support a three-component framework of WM (Oberauer, 2002; 2005; 2009; Oberauer & Hein, 2012), which proposes that an intermediary "region of direct access" can maintain items outside of focal attention, but in an accessible state, ready to be used for ongoing processing. In the current double-retrocue WM task, the latent item would be retained in the region of direct access while it is still relevant on the trial so that the participant can return it back to the focus of attention if it is subsequently cued, rather than reactivating the item from LTM. Congruent with this view, there was a decrement to WM of retaining the item outside focal attention, but returning it to focal attention did not affect subsequent LTM.

If retaining a latent item in WM is not accomplished by either continuous, active processing (e.g., rehearsal) of the item in focal attention or retrieving it from LTM, then what cognitive processes support their retention? Here, we consider two possibilities from the WM literature: refreshing and removal. First, refreshing is a hypothesized maintenance process that is proposed to be distinct from both rehearsal and episodic retrieval (Camos & Barrouillet, 2011; Johnson, 1992). Recently processed items that are outside of the current focus of attention, but still relevant for ongoing cognition (e.g., intermediary solutions in multi-step math problems), may be periodically refreshed by bringing them back into the focus of attention to retain them until they are no longer needed (see Camos et al., 2018 for review). Thus, refreshing could provide a cognitive mechanism for the observed pattern of results that switching attention impacted WM, but had no impact on LTM in any of the groups. However, there are at least two theoretical assumptions about how refreshing operates that conflict with this possibility. First, dividing attention with a secondary task is thought to disrupt refreshing (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). If refreshing is required to reactivate latent items, then disrupting attention with the parity task in Experiment 2 should have resulted in a stronger difference between probe 2 switch versus both probe 1 and probe 2 stay trials in young adults with DA compared to young adults with FA in Experiment 1. Although DA impaired performance overall, both young adult groups showed a similar decrease between probe 2 switch and both probe 1 and probe 2 stay trials. This indicates that disrupting attention did not disproportionately impact the reactivation of latent items back into focal attention. A second theoretical assumption regarding refreshing held by some researchers is that its use in WM should impact subsequent retrieval from LTM (Camos & Portrat, 2015; Jarjat, Hoareau, Plancher, Hot, Lemaire, & Portrat, 2018; Johnson, Reeder, Raye, & Mitchell, 2002; but see Rose et al., 2014 for an alternate view). As already discussed previously, there was no impact of having reactivated latent (UA) items in WM on subsequent LTM compared to initially attended but dropped (AU) items, nor was there an impact of disrupting attention on LTM in Experiment 2. The current results also weigh in on theoretical assumptions regarding the ways that aging impairs refreshing ability (Johnson et al., 2002; Loaiza & McCabe, 2013). Although there were age effects on WM, as we discuss further in the next section, an alternative explanation can account for this finding, and there was no age effect (or interaction with cueing) on subsequent LTM. Overall, these results make refreshing an unlikely candidate cognitive mechanism for why switching attention affected WM, but not LTM.

Another possible cognitive mechanism underlying the current pattern of results may be removal, which is coherent within the aforementioned three-component framework of WM (Oberauer, 2002, 2009, 2019). Removal is the process by which access to currently-irrelevant information in WM is reduced in order to facilitate the maintenance of relevant information (see Lewis-Peacock et al., 2018 for review). Researchers have argued that removal may operate either temporarily (and thus reversibly) on the information itself or permanently by unbinding the information that is no longer relevant on the trial from its context in WM. Thus, in the current paradigm, latent items in WM may have been temporarily removed from WM, resulting in an overall switch cost that impacted WM; but removal of the item-context bindings that were relevant on a trial did not affect subsequent LTM, even when the outcome measure of LTM assessed memory for the binding of different contextual details of the original information that was maintained in WM. This would suggest that removal operates by temporarily reducing the accessibility of currently irrelevant item-context bindings rather than permanently affecting their consolidation. This hypothesis about a temporary removal process is consistent with results (e.g., TMS-induced reactivation effects) that have shown a difference between the reactivation of currently irrelevant information that may become relevant later on in a WM trial, but not for items after they were cued as no longer relevant on the trial (Fulvio & Postle, 2020; Rose et al., 2016; Wolff et al., 2017). Future research will be required to fully address this possibility, particularly with regards to whether the consequence of selectively attending to a cued item and the removal of the uncued item are two sides of the same coin, as suggested by the biased-competition model of selective attention (Duncan & Desimone, 1995). For example, recent fMRI decoding results of a variant of our two-item double-retrocue task in which the items remained on the screen during the delay--thereby turning this WM task into a selective attention task--are consistent with this view (Sheldon, Saad, Sahan, Meyering, Starrett, LaRocque, Rose, & Postle, 2021).

What neurobiological mechanisms might support the short-term retention of information that is neither actively retained in focal attention or passively retained in LTM? An activity-silent WM account, such as the Synaptic Theory of WM (Mongillo et al., 2008; Stokes, 2015; Trübutschek et al., 2017), suggests that short-term synaptic plasticity mechanisms might be responsible for retaining such items. Information about an item that is being actively processed in WM may be actively retained by sustained, elevated spiking of neurons coding for the item, but if the item is dropped from focal attention, the information may be temporarily stored in an intermediate state via short term synaptic-plasticity mechanisms, which can facilitate its reactivation if participants shift attention back to this deprioritized item. In this case, our data suggest that recovering deprioritized items might not need LTM retrieval; instead, deprioritized items may be stored in an intermediate memory state via short-term synaptic plasticity mechanisms that bridge the gap from WM to the beginning stages of consolidation in LTM.

How might the cognitive concepts of refreshing and removal be incorporated into this neurobiological account? The cognitive concepts of refreshing and removal may be represented by shifts in oscillatory patterns of functional connectivity between both top-down signals, from control regions in frontal and parietal cortex, and bottom-up signals, from the corresponding hemispheres in posterior, sensory cortex. Changes in functional connectivity between frontal and posterior regions in the left and right hemispheres may enable the transition between memory states for items presented on the left and right visual fields, as was done in the current task (Myers, Walther, Wallis, Stokes, & Nobre, 2015; Myers, Stokes, & Nobre, 2017). In this task, both the left and right item would be represented in focal attention via sustained activation, the shift to selectively attending to the cued item (e.g., the left item) while passively retaining the uncued (e.g., right item), and then the switch to reactivating the initially uncued item and removing the no longer relevant item from WM could be accomplished via interhemispheric oscillatory dynamics (Bonnefond & Jensen 2012; Foster, Sutterer, Serences, Vogel, & Awh, 2016). Cross-frequency coupling between frontal theta and posterior alpha/beta frequencies could support the selective attention of a cued item initially represented in right visual cortex and the simultaneous deprioritization of the uncued item represented in left visual cortex through interhemispheric shifts in alpha oscillations or patterns of functional connectivity between posterior representational regions and frontal-parietal control regions (Günseli, Fahrenfort, Van Moorselaar, Daoultzis, Meeter, & Olivers, 2019; Schneider, Göddertz, Haase, Hickey, & Wascher, 2019).

To summarize, the current results suggest that the strong effects of retrocueing, age, and dividing attention in WM had no impact on subsequent LTM. This replicates and extends the results of LaRocque and colleagues (2015) to suggest that reactivating latent items in WM does not require retrieval from LTM. Instead, the results cohere with an activity-silent account of WM, such that latent items are temporarily supported via short-term synaptic-plasticity mechanisms.

LTM deficits do not explain age differences in WM

A critical question emerging from the cognitive aging literature concerns whether older adults can shift items in and out of focal attention as well as young adults, particularly via retrocues. As discussed in the introduction, most prior research examining age effects on retrocue benefits has used tasks with a single retrocue. Here, our double-retrocue paradigm allowed us to investigate whether the age deficits in episodic LTM, commonly observed in the literature and in the current study, may explain any disproportionate switching cost in WM as well as a reduced benefit of reactivating latent items to subsequent LTM. That is, if reactivating latent WM items requires retrieval from LTM, then (1) older adults should show worse WM performance for probe 2 switch versus stay trials (pre-registered hypothesis 1), and (2) older adults should exhibit smaller benefits to LTM for reactivated latent (UA) items compared to younger adults (pre-registered hypothesis 2B).

Regarding the first prediction, we observed that older and younger adults showed a similar decrease to WM performance for probe 2 switch trials compared to probe 2 stay trials. This suggests that older adults were able to use retrocues in WM similarly to younger adults, consistent with prior work (Loaiza & Souza, 2018), and further indicates that reactivation of latent items in WM is unlikely to require LTM if older adults who tend to have LTM deficits can switch attention as well as younger adults. However, in our data, the way that the older adults used the retrocues appeared to differ from young adults. Relative to probe 1, younger adults' accuracy improved on probe 2 stay trials, which is sensible given that the same item was cued and tested on both probe 1 and probe 2 stay trials, and participants received feedback about their recognition decision on probe 1. Older adults did not benefit from this repeated testing with feedback. Instead, they showed a decrease in accuracy on probe 2 switch trials compared to both probe 1 and probe 2 stay trials, the latter of which yielded similar accuracy. This suggests that older adults were not fully following the retrocues to proactively attend to the cued item to prepare for the recognition probes; instead, they may have relied on reacting to the probes and making their match/nonmatch decision based on the strength of the familiarity signal. The results of an experiment that manipulated the familiarity of the stimuli supported this hypothesis (Xu, Chao, & Rose, 2022). Older adults' overreliance on a familiarity signal for their recognition decisions on the WM task thus may be a critical factor that contributed to their performance on this double-retrocue task. Overall, the current results suggest that older adults' WM deficit was not due to a requirement to retain the items in and retrieve them from LTM.

As explained previously, there were no differences among the retrocue conditions (AA, AU, UA, UU) on LTM for any of the age groups, thus conflicting with the second pre-registered prediction that reactivating latent items requires retrieval from LTM that in turn impacts subsequent LTM. This is interesting given the previous literature concerning retrocue effects on WM and LTM in young and older adults. Strunk, Morgan, Reaves, Verhaeghen & Duarte (2019) demonstrated a retrocue effect on a single retrocue WM task that had lasting effects on subsequent LTM for both young and older adults, but only for item memory--not location memory. One major difference between the current study and Strunk et al. (2019) is that the double-retrocue WM task in the current study allowed us to test the nature of retaining and reactivating latent WM. Despite the differences, we also observed a retrocue effect for both the young and old groups on the WM test. However, cueing was not found to influence subsequent item, location, or associative LTM. Perhaps retrocues function to enhance the temporary bindings of item-location associations in WM, but these temporary bindings are dissolved at the end of each trial (Oberauer, 2005). That is, they may support WM, but they are not strongly consolidated so that they would affect subsequent LTM. Another possibility is that the nature of the differences between the WM tasks caused the divergent results in LTM. Cued items are to be actively maintained in both single- and double-retrocue tasks, whereas initially uncued items may be passively retained in double-retrocue tasks. Differences in the nature of the maintenance processes in the WM tasks may be the source of differences in the pattern of LTM performance between the studies. The development of a Dynamic Processing Model of WM (Rose, 2020) is designed to help accommodate the variability in the effects across different contexts including different domains of WM (verbal, visual, spatial), stimuli (familiar, novel), tasks (recall, recognition), conditions (with vs. without distraction), and populations (young adults, older adults), etc.

Limitations and future directions

Many studies show that adults have the capacity to actively maintain and sustain up to three to four items in WM (Adam, Mance, Fukuda & Vogel, 2015; Luria, Balaban, Awh & Vogel, 2016; Vogel & Machizawa, 2004). In our experiments, however, two items in a single trial were used. It is possible that participants actively maintained both items in focal attention throughout the trial. If that is the case, then both attended and latent items may never have dropped to LTM; therefore, recovering latent items on this task may not involve retrieval from LTM. However, several studies with basically identical paradigms (Lewis-Peacock et al., 2012;

Larocque et al., 2013; Rose et al., 2016) have shown the return-to-baseline levels of decoding for an uncued item with only two memory items per trial, which suggests that participants do not maintain these items in WM in an active, sustained manner, but they can be reactivated by noninvasive brain stimulation or when attention is shifted back to them. Experiment 2 (DA) showed that when participants processed the secondary task, the latent items were likely dropped from focal attention due to the demanding nature of the DA task. It might be difficult for participants to actively maintain two items and successfully process the secondary task. Taken together, the latent item should be dropped from the focus of attention in our double-retrocue WM task.

The fact that older and younger adults performed similarly on the subsequent LTM test is also peculiar given the extensive literature showing such age differences, particularly in location and associative memory (Old & Naveh-Benjamin, 2008). At first glance the current sample of older adults could be considered “super-agers”, but we hasten to point out that they showed typical deficits on independent measures of LTM (i.e., the TICS initial and final free recall tests) compared to younger adults. Thus, we interpret the lack of age and cue effects on LTM as genuine rather than owing to a high-performing sample of older adults.

Examining the effects of actively vs. passively maintaining information in WM is important to elucidate the dynamics of WM and advancing WM theory. Because of the similarity between this double-retrocue task and that of previous studies showing evidence for “activity-silent” retention, as well as the performance differences between probe 2 switch versus stay trials, it is reasonable to infer that the uncued items were passively retained. Nevertheless, future studies should include neuroimaging, brain decoding and brain stimulation techniques to investigate the role of LTM in WM and the effects of retrocues, aging, and divided attention on the interaction between latent WM and LTM. Studying these topics can help this field to better understand the relationship between WM and LTM. We also note that including neutral-cue or invalid-cue trials could serve as useful control conditions to gauge the size of retrocue benefits as in previous studies (e.g., LaRocque et al., 2015; Strunk et al., 2019).

Conclusion

Our pre-registered report predicted that if retaining and reactivating latent items in WM involves episodic LTM processes then doing so would affect subsequent LTM. However, the results showed that, although the retrocue manipulation affected WM, as did the effects of age and divided attention, none of these factors had consistent, reliable effects on LTM. Thus, the results are inconsistent with the LTM-retrieval hypothesis that latent WM items are retrieved via LTM processes, at least for this double-retrocue task with only two items to remember. Instead, latent items may be retained on such tasks via short-term synaptic plasticity mechanisms as predicted by the activity-silent WM account (Mongillo et al., 2008; Stokes, 2015; Trübutschek et al., 2017). The cognitive concepts described as the region of direct access (Oberauer, 2002, 2009, 2019) and removal (Lewis-Peacock et al., 2018) may refer to these neurobiological retention mechanisms.

One's ability to briefly drop information from focal attention and get it back again, as in two-factor authentication of online accounts, represents the beginning stages of the transition from active processing in WM to passive, latent representation in LTM. This ability is one of the most intriguing areas of research on WM and LTM, yet it remains relatively poorly understood. However, cognitive, neural, and neurocomputational models of WM are increasingly incorporating mechanisms that attempt to account for this interaction with LTM. We hope that the results reported here help to further refine such models as they attempt to detail the precise nature of interactions among attention, WM, and LTM processes.

Author Note

This research was supported by the National Science Foundation's CAREER Grant #1848440 awarded to N.S.R. This study's design and hypotheses were pre-registered; all data have been made publicly available and can be accessed at https://osf.io/z9cgq/. Author Contributions: Chang-Mao Chao: data curation, formal analysis, methodology, software, investigation, visualization, writing (original draft, review, and editing); Chenglinxi Xu: data curation, investigation, writing (review & editing); Vanessa Loaiza: writing (review & editing); Nathan Rose: conceptualization, funding acquisition, methodology, project administration, resources, software, supervision, writing (original draft, review, and editing). The authors would also like to acknowledge Lauren Crowe, Cheick Diallo, Geralynn Timm, Bernadette Widjawa, and Yuki Yoshioka for their assistance with data collection, and Joseph Saito, Morgan Widhalm, and Nicholas Yeh for their help with an initial pilot study.

References

Adam, K. C., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The contribution of attentional lapses to individual differences in visual working memory capacity. *Journal of cognitive neuroscience*, *27*(8), 1601-1616.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of communication disorders*, *36*(3), 189-208.

Baddeley, A. (2012). Working memory: theories, models, and controversies. Annual review of psychology, 63, 1-29.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*(1), 83.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *33*(3), 570.

Basak, C., & Verhaeghen, P. (2011). Aging and switching the focus of attention in working memory: age differences in item availability but not in item accessibility. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66(5), 519-526.

Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In Robert L. Solso (ed.), Information Processing and Cognition: The Loyola Symposium. Lawrence Erlbaum. pp. 123--144.

Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. Practical aspects of memory: Current research and issues, 1, 396-401.

Bonnefond M, Jensen O. 2012. Alpha oscillations serve to protect working memory maintenance against anticipated distracters. Current Biology 22(20):1969–1974. DOI 10.1016/j.cub.2012.08.029.

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

Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences*, *108*(27), 11252-11255.

Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: From passive maintenance to active refreshing. Developmental psychology, 47(3), 898.

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory?: What is attentional refreshing?. *Annals of the new York Academy of Sciences*, *1424*(1).

Camos V., Lagner P., Barrouillet P. (2009). Two maintenance mechanisms of verbal information in working memory. *J. Mem. Lang*. 61 457–469 10.1016/j.jml.2009.06.002

Camos, V., & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic bulletin & review*, *22*(4), 1029-1034.

Constantinidis, C., Funahashi, S., Lee, D., Murray, J. D., Qi, X. L., Wang, M., & Arnsten, A. F. (2018). Persistent spiking activity underlies working memory. *Journal of neuroscience*, *38*(32), 7020-7028.

Cowan, N. (1995). Attention and memory: An integrated framework New York: Oxford University Press.[Google Scholar].

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, *24*(1), 87-114.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. Progress in brain research, 169, 323-338.

Castel, A. D., & Craik, F. I. M. (2003). The effects of aging and divided attention on memory for item and associative information. Psychology and Aging, 18, 873–885. doi:10.1037/0882-7974.18.4.873

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. The Behavioral and Brain Sciences, 24(1), 87–114; discussion 114-85. https://doi.org/10.1017/S0140525X01003922

Craik, F.I.M. (1970). Fate of primary memory items in free recall. Journal of Verbal Learning & Verbal Behavior, 9, 143-148.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short- term memory. Journal of Verbal Learning and Verbal Behavior, 12,. 599–607.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In Craik, F. I. M., & S. E. Trehub (Eds.), Aging and cognitive processes (pp. 191–211). New York: Plenum Press.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*(1), 193-222.

Duarte, A., Hearons, P., Jiang, Y., Delvin, M. C., Newsome, R. N., & Verhaeghen, P. (2013). Retrospective attention enhances visual working memory in the young but not the old: An ERP study. *Psychophysiology*, *50*(5), 465-476.

Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *Journal of Neuroscience*, *33*(15), 6516-6523.

Fiebig, F., & Lansner, A. (2017). A spiking working memory model based on Hebbian short-term potentiation. *Journal of Neuroscience*, *37*(1), 83-96.

Foster, Vogel, and Awh. In press. Working memory as persistent neural activity. To appear in: M.J. Kahana and A.D. Wagner (Eds.), Oxford Handbook of Human Memory. Oxford University Press.

Foster JJ, Sutterer DW, Serences JT, Vogel EK, Awh E. 2016. The topography of alpha-band activity tracks the content of spatial working memory. Journal of Neurophysiology, 115(1):168–177 DOI 10.1152/jn.00860.2015

Fulvio, J. M., & Postle, B. R. (2020). Cognitive Control, Not Time, Determines the Status of Items in Working Memory. *Journal of Cognition*, 3(1), 8. DOI: http://doi.org/10.5334/joc.98

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*(3997), 652-654.

Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: Where do they come from?. *Journal of experimental child psychology*, *110*(3), 469-479.

Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., & D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proceedings of the National Academy of Sciences*, *105*(35), 13122-13126

Gilchrist, A. L., Duarte, A., & Verhaeghen, P. (2016). Retrospective cues based on object features improve visual working memory performance in older adults. *Aging*, *Neuropsychology*, *and Cognition*, *23*(2), 184-195.

Günseli E, Fahrenfort JJ, Van Moorselaar D, Daoultzis KC, Meeter M, Olivers CNL. 2019. EEG dynamics reveal a dissociation between storage and selective attention within working memory. Scientific Reports 9(1):13499 DOI 10.1038/s41598-019-49577-0.

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.

Jarjat, G., Hoareau, V., Plancher, G., Hot, P., Lemaire, B., & Portrat, S. (2018). What makes working memory traces stable over time?. *Annals of the New York Academy of Sciences*, *1424*(1), 149-160.

Jennings, J. M., & Jacoby, L. L. (2003). Improving memory in older adults: Training recollection. *Neuropsychological Rehabilitation*, 13(4), 417-440.

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of cognitive neuroscience*, *4*(3), 268-280.

Johnson, M. K., Reeder, J. A., Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychological Science*, *13*(1), 64-67

Johnson, W., Logie, R. H., & Brockmole, J. R. (2010). Working memory tasks differ in factor structure across age cohorts: Implications for dedifferentiation. *Intelligence*, *38*(5), 513-528.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. Neuroscience, 139(1), 181-193.

Jonides, J., Marshuetz, C., Smith, E. E., Reuter-Lorenz, P. A., Koeppe, R. A., & Hartley, A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. Journal of Cognitive Neuroscience, 12, 188 –196. doi:10.1162/089892900561823

Jost, K., Bryck, R. L., Vogel, E. K., & Mayr, U. (2011). Are old adults just like low working memory young adults? Filtering efficiency and age differences in visual working memory. *Cerebral cortex*, *21*(5), 1147-1154.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. science, 319(5865), 966-968.

Knopman, D. S., Roberts, R. O., Geda, Y. E., Pankratz, V. S., Christianson, T. J., Petersen, R. C., & Rocca, W. A. (2010). Validation of the telephone interview for cognitive status-modified in subjects with normal cognition, mild cognitive impairment, or dementia. *Neuroepidemiology*, 34(1), 34-42.

LaRocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in human neuroscience*, *8*, 5.

LaRocque, J. J., Eichenbaum, A. S., Starrett, M. J., Rose, N. S., Emrich, S. M., & Postle, B. R. (2015). The short-and long-term fates of memory items retained outside the focus of attention. Memory & Cognition, 43(3), 453-468.

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of cognitive neuroscience*, *24*(1), 61-79.

Loaiza, V. M., McCabe, D. P., Youngblood, J. L., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(5), 1258.

Loaiza, V. M., & McCabe, D. P. (2013). The influence of aging on attentional refreshing and articulatory rehearsal during working memory on later episodic memory performance. *Aging*, *Neuropsychology*, *and Cognition*, *20*(4), 471-493.

Loaiza, V. M., & Camos, V. (2016). Does controlling for temporal parameters change the levels-of-processing effect in working memory?. *Advances in Cognitive Psychology*, *12*(1), 2.

Loaiza, V. M., & Souza, A. S. (2018). Is refreshing in working memory impaired in older age? Evidence from the retro-cue paradigm. *Annals of the New York Academy of Sciences*, *1424*(1), 175-189.

Loaiza, V. M., & Halse, S. C. (2019). Where working memory meets long-term memory: The interplay of list length and distractors on memory performance. Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(8), 1455.

Loaiza, V. M., & Souza, A. S. (2019). An age-related deficit in preserving the benefits of attention in working memory. *Psychology and Aging*, *34*(2), 282.

Loaiza, V. M., Doherty, C., & Howlett, P. (2021). The long-term consequences of retrieval demands during working memory. Memory & cognition, 49(1), 112-126.

Loaiza, V. M., & Lavilla, E. T. (2021). Elaborative strategies contribute to the long-term benefits of time in working memory. Journal of Memory and Language, 117, 104205.

Loosli, S. V., Rahm, B., Unterrainer, J. M., Weiller, C., & Kaller, C. P. (2014). Developmental change in proactive interference across the life span: Evidence from two working memory tasks. Developmental psychology, 50(4), 1060.

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100-108.

McCabe, J., & Hartman, M. (2008). Working memory for item and temporal information in younger and older adults. Aging, Neuropsychology, and Cognition, 15, 574 – 600. doi:10.1080/13825580801956217

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, *58*(2), 480-494.

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222.

McElree, B. (1998). Attended and non-attended states in working memory: Accessing categorized structures. *Journal of Memory and Language*, *38*(2), 225-252.

McElree, B. (2006). Accessing recent events. Psychology of Learning and Motivation - Advances in Research and Theory, 46(06), 155–200. https://doi.org/10.1016/S0079-7421(06)46005-9

McNab, F., Zeidman, P., Rutledge, R. B., Smittenaar, P., Brown, H. R., Adams, R. A., & Dolan, R. J. (2015). Age-related changes in working memory and the ability to ignore distraction. *Proceedings of the National Academy of Sciences*, *112*(20), 6515-6518.

Mok, R. M., Myers, N. E., Wallis, G., & Nobre, A. C. (2016). Behavioral and neural markers of flexible attention over working memory in aging. *Cerebral Cortex*, *26*(4), 1831-1842.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, *319*(5869), 1543-1546.

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing information during working memory: beyond sustained internal attention. *Trends in cognitive sciences*, 21(6), 449-461.

Myers NE, Walther L, Wallis G, Stokes MG, Nobre AC. 2015. Temporal dynamics of attention during encoding versus maintenance of working memory: complementary views from event-related potentials and alpha-band oscillations. Journal of Cognitive Neuroscience, 27(3):492–508 DOI 10.1162/jocn_a_00727.

Newsome, R. N., Duarte, A., Pun, C., Smith, V. M., Ferber, S., & Barense, M. D. (2015). A retroactive spatial cue improved VSTM capacity in mild cognitive impairment and medial temporal lobe amnesia but not in healthy older adults. *Neuropsychologia*, *77*, 148-157.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(3), 411–421. https://doi.org/10.1037//0278-7393.28.3.411

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.

Oberauer, K. (2009). Design for a working memory. *Psychology of learning and motivation*, *51*, 45-100.

Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, *21*(3), 164-169.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., ... & Ward, G. (2018). Benchmarks for models of short-term and working memory. Psychological bulletin, 144(9), 885.

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.

Park, D. C. (2000). The basic mechanisms accounting for age-related decline in cognitive function. Cognitive aging: A primer, 11(1), 3-19.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* 10.3758/s13428-018-01193-y

Schmiedek, F., Li, S.-C., & Lindenberger, U. (2009). Interference and facilitation in spatial working memory: Age-associated differences in lure effects in the n-back paradigm. Psychology and Aging, 24, 203–210. doi:10.1037/a0014685

Schneider D, Göddertz A, Haase H, Hickey C, Wascher E. 2019. Hemispheric asymmetries in EEG alpha oscillations indicate active inhibition during attentional orienting within working memory. *Behavioural Brain Research* 359:38–46 DOI 10.1016/j.bbr.2018.10.020.

Sheldon, A.D., Saad, E., Sahan, M.I., Meyering, E., Starrett, M.J., LaRocque, J.J., Rose, N.S., & Postle, B.R. (2021). Attention biases competition for visual representation via enhancement of targets and inhibition of nontargets. *Journal of Cognitive Neuroscience*.

Singh, B., Gambrell, A., & Correll, J. (2022). Face Templates for the Chicago Face Database. *Behavior Research Methods*, 1-7.

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1237.

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 20-31.

Souza, A. S. (2016). No age deficits in the ability to use attention to improve visual working memory. Psychology and Aging, 31(5), 456.

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retrocue effect. *Attention*, *Perception*, *& Psychophysics*, *78*(7), 1839-1860.

Souza, A. S., & Oberauer, K. (2017). Time to process information in working memory improves episodic memory. Journal of Memory and Language, 96, 155-167.

Stokes, M. G. (2015). 'Activity-silent'working memory in prefrontal cortex: a dynamic coding framework. *Trends in cognitive sciences*, *19*(7), 394-405.

Strunk, J., Morgan, L., Reaves, S., Verhaeghen, P., & Duarte, A. (2019). Retrospective attention in short-term memory has a lasting effect on long-term memory across age. *The Journals of Gerontology: Series B*, *74*(8), 1317-1325.

Trübutschek, D., Marti, S., Ojeda, A., King, J. R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *Elife*, *6*, e23871.

Verhaeghen, P., & Basak, C. (2005). Ageing and switching of the focus of attention in working memory: Results from a modified N-Back task. The Quarterly Journal of Experimental Psychology Section A, 58(1), 134-154.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748-751.

Rose, N. S., Myerson, J., Roediger III, H. L., & Hale, S. (2010). Similarities and differences between working memory and long-term memory: evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 471.

Rose, N. S., & Craik, F. I. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1019.

Rose, N. S., Buchsbaum, B. R., & Craik, F. I. (2014). Short-term retention of a single word relies on retrieval from long-term memory when both rehearsal and refreshing are disrupted. *Memory & Cognition*, *42*(5), 689-700.

Rose, N. S., Craik, F. I., & Buchsbaum, B. R. (2015). Levels of processing in working memory: Differential involvement of frontotemporal networks. *Journal of cognitive neuroscience*, *27*(3), 522-532.

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136-1139.

Rose, N. S. (2020). The dynamic-processing model of working memory. *Current Directions in Psychological Science*, *29*(4), 378-387.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. Psychological review, 114(1), 104.

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature neuroscience*, 20(6), 864-871.

Yi, Y., & Friedman, D. (2014). Age-related differences in working memory: ERPs reveal age-related delays in selection-and inhibition-related processes. *Aging, Neuropsychology, and Cognition*, *21*(4), 483-513.

Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1363-1374.

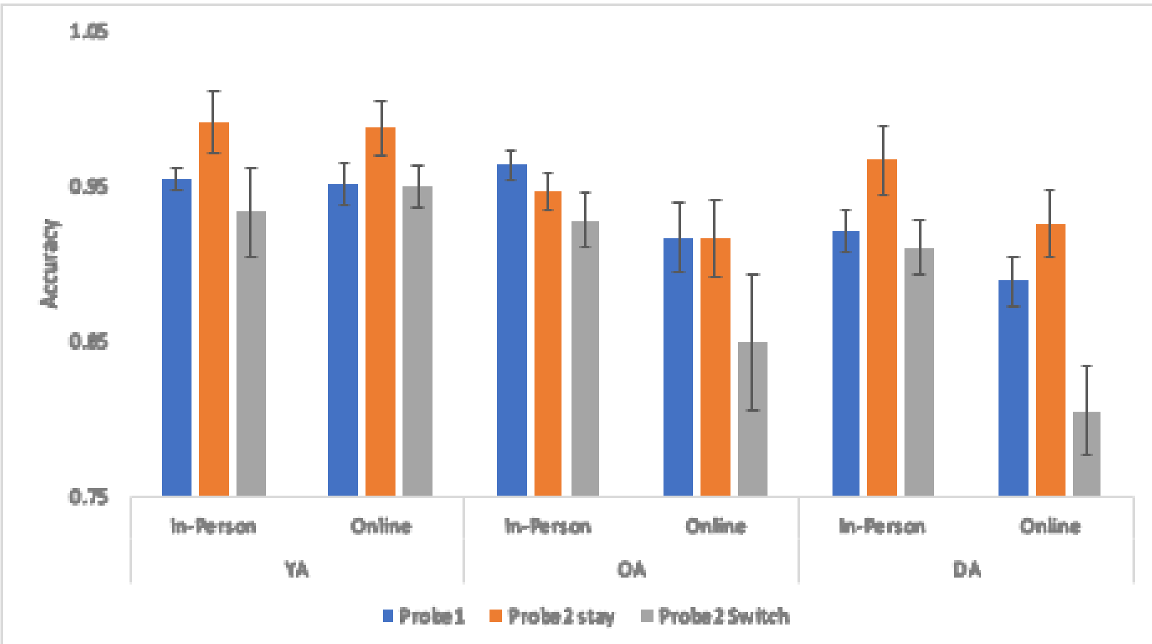
Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2017). Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, *40*(6), 1452-1464.

Supplementary Material

Online vs. In-person testing

As discussed in the text, participants completed all tasks either online on their personal computer in a distraction free environment or in-person, in our laboratory on a computer in a distraction-free testing station. To confirm that the location of testing did not interact with any of our key factors of interest, ANOVAs were conducted on the data separated by testing location. The means and results are reported below. Although performance was somewhat poorer in the online testing procedures, none of the interactions were significant.

WM analyses separated by location of testing (online vs. in-person)



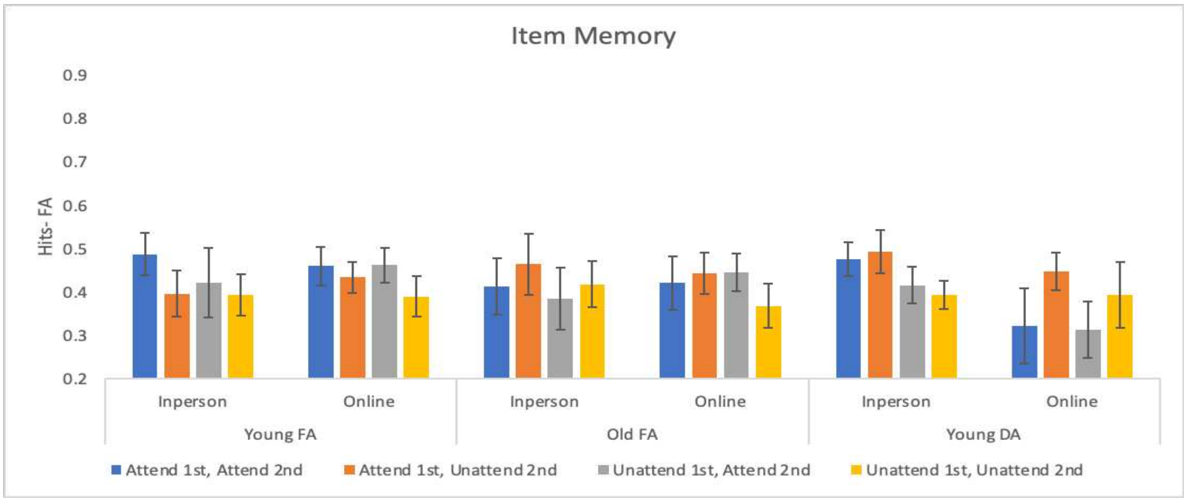
Supplemental Figure 1: WM analyses separated by location of testing (online vs. in-person)

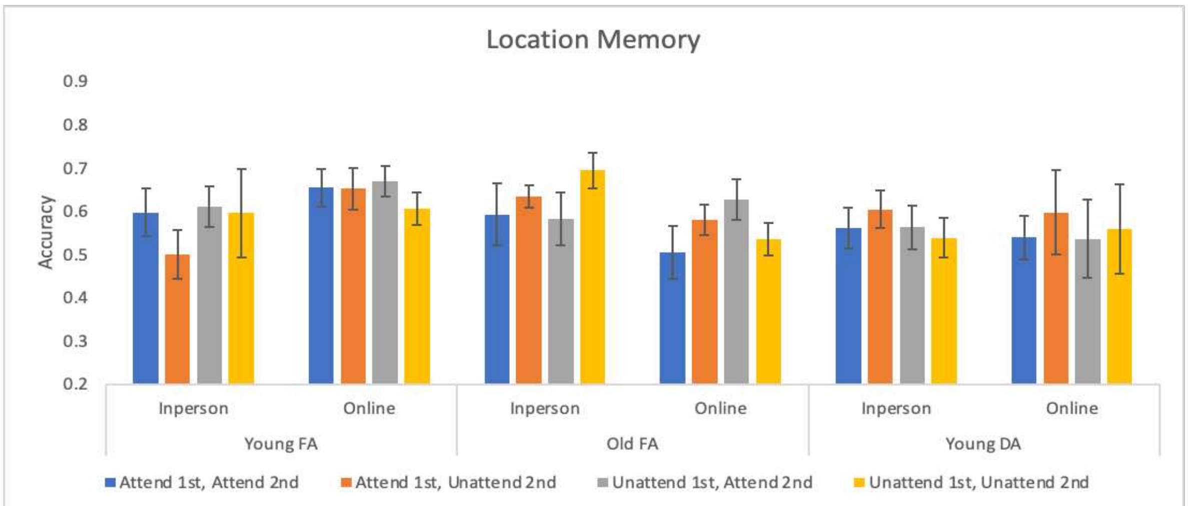
For YA (In Person vs. Online): Probe-type: $F(2,56) = 13.04$, $p < 0.001$, $\eta p2 = 0.318$
In Person/Online: $F(2, 56) = 0.71$, $p = 0.495$, $\eta p2 = 0.025$
Probe-type*In Person/Online: $F(1,28) = 0.04$, $p = 0.849$, $\eta p2 = 0.001$

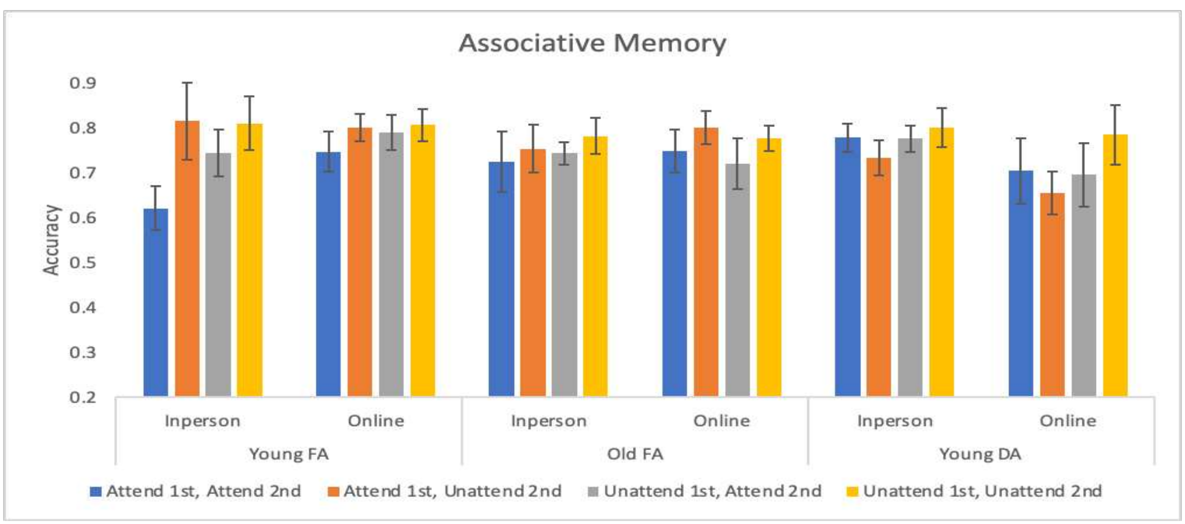
For OA (In Person vs. Online): Probe-type: $F(2,50) = 7.58$, $p = 0.001$, $\eta p2 = 0.233$
In Person/Online: $F(2, 50) = 4.44$, $p = 0.045$, $\eta p2 = 0.151$
Probe-type*In Person/Online: $F(1,25) = 1.50$, $p = 0.233$, $\eta p2 = 0.057$

For DA (In Person vs. Online): Probe-type: $F(2,54) = 14.98$, $p < 0.001$, $\eta p2 = 0.357$
In Person/Online: $F(2, 54) = 9.18$, $p = 0.005$, $\eta p2 = 0.254$
Probe-type*In Person/Online: $F(1,27) = 3.04$, $p = 0.056$, $\eta p2 = 0.101$

LTM analyses separated by location of testing (online vs. in-person)







For YA (In Person vs. Online):

Item memory: condition: $F(3,84) = 2.45$, $p = 0.07$, $\eta p2 = 0.081$

In Person/Online: $F(1, 28)= 0.04$, $p = 0.837$, $\eta p2 = 0.002$

	condition*In Person/Online: $F(3,84) = 0.54$, $p = 0.653$, $\eta_p^2 = 0.019$
Location memory:	condition: $F(3,84) = 0.73$, $p = 0.54$, $\eta_p^2 = 0.025$
	In Person/Online: $F(1, 28) = 2.23$, $p = 0.142$, $\eta_p^2 = 0.076$
	condition*In Person/Online: $F(3,84) = 0.82$, $p = 0.485$, $\eta_p^2 = 0.029$
Associative memory:	condition: $F(3,84) = 3.39$, $p = 0.022$, $\eta_p^2 = 0.108$
	In Person/Online: $F(1, 28) = 0.98$, $p = 0.331$, $\eta_p^2 = 0.034$
	condition*In Person/Online: $F(3,84) = 1.00$, $p = 0.395$, $\eta_p^2 = 0.035$

For OA (In Person vs. Online):

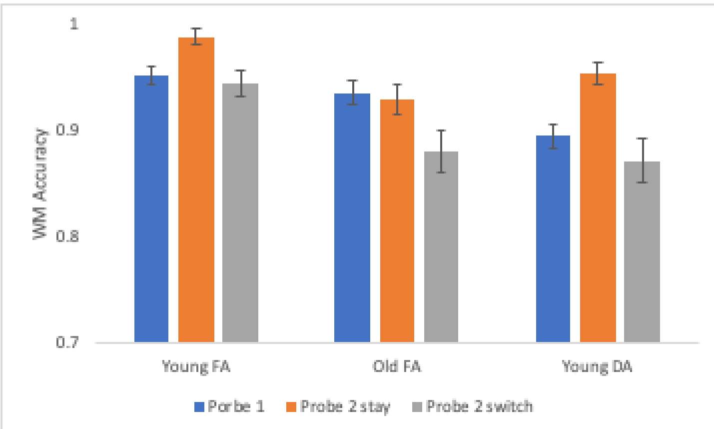
Item memory:	condition: $F(3,75) = 1.25$, $p = 0.298$, $\eta_p^2 = 0.048$
	In Person/Online: $F(1, 25) < 0.001$, $p = 0.992$, $\eta_p^2 < 0.001$
	condition*In Person/Online: $F(3,75) = 1.09$, $p = 0.359$, $\eta_p^2 = 0.042$
Location memory:	condition: $F(3,75) = 0.99$, $p = 0.403$, $\eta_p^2 = 0.038$
	In Person/Online: $F(1, 25) = 2.18$, $p = 0.152$, $\eta_p^2 = 0.080$
	condition*In Person/Online: $F(3,75) = 1.93$, $p = 0.132$, $\eta_p^2 = 0.072$
Associative memory:	condition: $F(3,75) = 0.68$, $p = 0.565$, $\eta_p^2 = 0.027$
	In Person/Online: $F(1, 25) = 0.09$, $p = 0.763$, $\eta_p^2 = 0.004$
	condition*In Person/Online: $F(3,75) = 0.26$, $p = 0.856$, $\eta_p^2 = 0.010$

For DA (In Person vs. Online):

Item memory:	condition: $F(3,81) = 2.31$, $p = 0.082$, $\eta_p^2 = 0.079$
	In Person/Online: $F(1, 27) = 1.99$, $p = 0.169$, $\eta_p^2 = 0.069$
	condition*In Person/Online: $F(3,81) = 1.28$, $p = 0.288$, $\eta_p^2 = 0.045$
Location memory:	condition: $F(3,81) = 0.34$, $p = 0.797$, $\eta_p^2 = 0.012$
	In Person/Online: $F(1, 27) = 0.05$, $p = 0.83$, $\eta_p^2 = 0.002$
	condition*In Person/Online: $F(3,81) = 0.05$, $p = 0.984$, $\eta_p^2 = 0.029$
Associative memory:	condition: $F(3,81) = 1.78$, $p = 0.157$, $\eta_p^2 = 0.062$
	In Person/Online: $F(1, 27) = 2.22$, $p = 0.148$, $\eta_p^2 = 0.076$
	condition*In Person/Online: $F(3,81) = 0.26$, $p = 0.851$, $\eta_p^2 = 0.01$

WM analysis excluding the two young DA participants with 3 digits to verify during the delays

As discussed in the text, two young adults in the DA experiment had response times on the odd-even digit pretesting titration procedure that required them to respond to 3 rather than 2 digits during the maintenance and cueing phases of the WM test trials. The data were reanalyzed excluding these two subjects to confirm that the pattern of results would not change. The means and the ANOVA results comparing the young and old FA groups to the young DA participants who had only two digits to verify during each cue/delay period of the WM task are presented below. As may be seen, the same general pattern was observed.



Supplemental Figure 6 WM analysis (young DA with 3 digits are excluded)

Young FA vs Young DA: Probe Type: $F(2,110) = 26.53$, $p < 0.001$, $\eta p2 = 0.325$
Group: $F(1,55) = 17.16$, $p < 0.001$, $\eta p2 = 0.238$
Probe Type*Group $(2,110) = 2.16$, $p = 0.121$, $\eta p2 = 0.038$

Old FA vs Young DA: Probe Type: $F(2,104) = 18.56$, $p < 0.001$, $\eta p2 = 0.263$
Group: $F(1,52) = 0.25$, $p = 0.622$, $\eta p2 = 0.005$
Probe Type*Group $(2,104) = 4.69$, $p = 0.011$, $\eta p2 = 0.083$

Response Time Analyses

As discussed in the text, accuracy was the main dependent variable of interest. Median response times (RTs) are reported here for the interested reader.

Supplemental Table 1. Median (SEM) RTs on correct trials for Young and Older Adult groups with Full Attention and the Young Adult group with Divided Attention.

Median (SEM)	Probe1	Probe2 Stay	Probe2 Switch
Young FA	0.670(0.006)	0.572(0.011)	0.671(0.008)
Old FA	0.902(0.008)	0.838(0.017)	0.921(0.010)
Young DA	1.016(0.007)	0.900(0.015)	1.090(0.011)

Experiment 1. Analysis of young and older adults' mean response times for the different trial types generally converged with interpretation of the accuracy data. Note that because the RT distributions were skewed, analyses were conducted on the log-transformed median for each condition and participant. An ANOVA of probe-type (probe 1, probe 2stay, probe 2 switch) by group (Young vs. Old) on reaction time revealed main effects of both probe-type [$F(2,110) = 45.93$, $p < 0.001$, $\eta p2 = 0.46$, BF_{10}=1.006e +12], and age group [$F(1,55) = 78.01$, $p < 0.001$, $\eta p2 = 0.59$, BF_{10}= 1.435e + 9]. Response times were faster for young than older adults. As can be seen in Supplemental Table 1, the interaction between probe-type and group was not significant [$F(2,110) = 2.09$, $p = 0.129$, $\eta p2 = 0.04$, BF_{01}= 2.12]. Like the accuracy data, the response time data suggest that both groups obtained a retrocue benefit.

Experiment 2. An ANOVA on probe-type (probe 1, probe 2 stay, probe 2 switch) by group (Young DA vs. Old FA) on reaction time revealed main effects of both probe-type [$F(2,108) = 22.06$, $p < 0.001$, $\eta p2 = 0.29$, BF_{10} = 141873], and age group [$F(1,54) = 8.32$, $p = 0.006$, $\eta p2 = 0.13$, BF_{10} = 7.856]. As may be seen in Supplemental Table 1, the interaction between probe-type and group was significant [$F(2,108) = 7.54$, $p < 0.001$, $\eta p2 = 0.12$, BF_{10} = 40.36]. For the old FA group, reaction time of probe 2 stay is faster than probe 1 [$t(26) = 4.24$, p

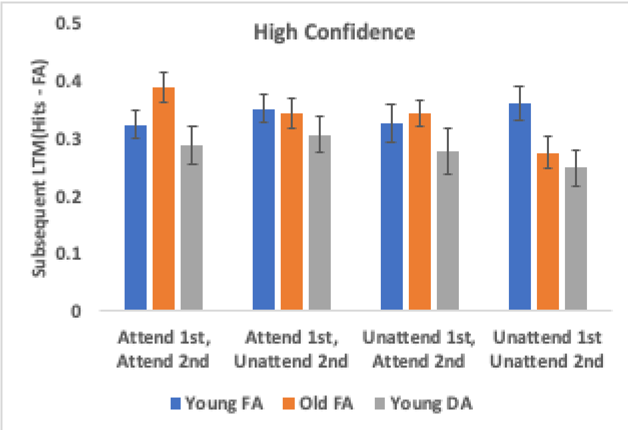
< 0.001, BF_{10} = 87.99] and probe 2 switch [t(26) = 5.27, p < 0.001, BF_{10} = 1285.41]. However, these effects were not found in the young DA group (probe 2 stay - probe 1:[t(28)= 0.7, p = 0.489, BF_{01} = 3.99]; probe 2 stay - switch:[t(28) = 2.03, p = 0.054, BF_{01} = 0.98]). A similar exploratory analysis was conducted to compare the young FA and young DA group. The effects of group [F(1,56) = 27.84, p < 0.001, ηp2 = 0.3, BF_{10} = 6782.01] and probe-type [F(2,112) = 20.28, p < 0.001, ηp2 = 0.27, BF_{10} = 209026.3] were significant. The two factors interacted with each other [F(2,112) = 7.99, p < 0.001, ηp2 = 0.13, BF_{10} = 74.1]. Simple main effect analysis showed that reaction time of probe 2 stay was faster than probe 1 [t(29) = 7.212, p < 0.001, BF_{10} = 446994.4] and probe 2 switch [t(29) = 5.20, p < 0.001, BF_{10} = 1795.02] for young FA.

LTM analyses separated by confidence

Although our preregistered hypotheses did not specify any a priori predictions about differences in the effects of retrocueing, age, or divided attention on subsequent LTM as a function of confidence, we report these exploratory analyses on recognition separated by high and low confidence to help guide future research for the interested reader. An important caveat with these analyses is that there may be differences in the numbers of high and low confidence judgments provided for items by each group for each condition. For example, the only comparison in which there was a group by condition interaction that differed between high and low confidence responses was on item recognition between young and older adults (F(3,165) = 3.47, p = 0.018, ηp2 = 0.059; all other ps>.05). The average hit-FA rate for high confidence judgments was higher for young than older adults for items that were unattended (i.e., untested) during both the first and second half of the trial (see Supplemental Figure 4); however, interpreting this age difference is complicated by the fact that young and older adults differed in the number of recognition decisions that were assigned with high confidence. The mean number of new items assigned high confidence (i.e., the FA rate) for young and older adults was 2.6 and 6.6 respectively, while the respective means for UU items were 3.9 and 6.0. An age difference in the average hit-FA rate for AA items also contributed to the significant interaction. This was because the average hit-FA rate for AA items assigned high confidence was greater for older than young adults (mean number of high confidence AA items was 5.2 and 5.5; note that the same within-subjects FA rate of 2.6 and 6.6 for young and older adults, respectively, applies here as well). There were no age differences in mean hit-FA rates between the critical AU or UA items. As discussed above, there were also no other group by condition interactions involving confidence for either location or associative LTM. Therefore, the main conclusions regarding LTM in this paper largely do not need to be qualified by these analyses separated by the level of confidence.

Supplemental Table 2. Mean (SEM) long-term recognition memory accuracy as a function of high (Absolutely) and log (Probably) confidence for Item Memory (top; hits-FAs), Location Memory (middle; hits), and Associative Memory (bottom; hits) for the Young and Older Adult groups with Full Attention and the Young Adult group with Divided Attention.

Item Memory	Absolutely				Probably			
Condition	AA	AU	UA	UU	AA	AU	UA	UU
Young FA	0.389(0.036)	0.345(0.027)	0.345(0.034)	0.275(0.029)	0.018(0.023)	0.048(0.023)	0.071(0.032)	0.115(0.028)
Old FA	0.442(0.046)	0.479(0.037)	0.446(0.038)	0.492(0.031)	0.044(0.025)	0.076(0.026)	0.013(0.023)	0.043(0.026)
Young DA	0.294(0.034)	0.315(0.031)	0.292(0.040)	0.254(0.032)	0.005(0.028)	0.085(0.036)	0.028(0.026)	0.140(0.028)
Location Memory								
Young FA	0.200(0.026)	0.230(0.026)	0.252(0.033)	0.225(0.020)	0.164(0.019)	0.233(0.029)	0.300(0.031)	0.225(0.026)
Old FA	0.151(0.027)	0.222(0.027)	0.148(0.031)	0.343(0.043)	0.173(0.028)	0.299(0.028)	0.241(0.042)	0.364(0.041)
Young DA	0.112(0.023)	0.121(0.018)	0.124(0.025)	0.092(0.020)	0.184(0.020)	0.253(0.034)	0.187(0.024)	0.247(0.024)
Associative Memory								
Young FA	0.178(0.019)	0.264(0.027)	0.275(0.030)	0.253(0.029)	0.133(0.020)	0.269(0.025)	0.281(0.031)	0.253(0.031)
Old FA	0.182(0.027)	0.247(0.026)	0.203(0.031)	0.308(0.036)	0.191(0.028)	0.275(0.037)	0.204(0.036)	0.389(0.042)
Young DA	0.224(0.028)	0.259(0.032)	0.218(0.034)	0.290(0.033)	0.178(0.021)	0.195(0.025)	0.198(0.027)	0.207(0.025)



Supplemental Figure 4. Mean hit-FA rates for high confidence item recognition decisions for young and older adults with full attention (FA) and young adults with divided attention (DA) as a function of condition.

The results of all main effects and interactions from the ANOVA:

FA vs. OA: condition: $F(3,165) = 0.97$, $p = 0.407$, $\eta p2 = 0.017$
Group: $F(1,55) = 0.005$, $p = 0.942$, $\eta p2 < 0.001$
condition*Group: $F(3,165) = 3.47$, $p = 0.018$, $\eta p2 = 0.059$

OA vs. DA: condition: $F(3,162) = 0.48$, $p = 0.699$, $\eta p2 = 0.009$
Group: $F(1,54) = 2.22$, $p = 0.142$, $\eta p2 = 0.04$
condition*Group: $F(3,162) = 0.96$, $p = 0.964$, $\eta p2 = 0.018$

YA vs. DA: condition: $F(3,171) = 4.23$, $p = 0.006$, $\eta p2 = 0.069$
Group: $F(1,57) = 2.45$, $p = 0.123$, $\eta p2 = 0.041$
condition*Group: $F(3,171) = 1.01$, $p = 0.390$, $\eta p2 = 0.017$