Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–24
© Experimental Psychology Society 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17470218231217723
qjep.sagepub.com
Sage

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

Chang Mao Chao[1], Chenlingxi Xu[1], Vanessa Loaiza[2] and Nathan S Rose[1]

Abstract

Switching one's focus of attention between to-be-remembered items in working memory (WM) is critical for cognition, but the mechanisms by which this is accomplished are unclear. A long-term memory (LTM) account suggests that switching attention away from an item, and passively retaining and reactivating such "latent" items back into the focus of attention involves episodic LTM retrieval processes, even for delays of only a few seconds. We tested this hypothesis using a two-item, double-retrocue WM task that requires participants to switch attention away from and reactivate items followed by subsequent LTM tests for reactivated items from the initial WM task (vs. continuously retained or untested control items). 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. The effects of reactivating latent items, as well as ageing and divided attention, had significant effects on WM performance, but did not interact with or systematically affect subsequent LTM for reactivated versus control items on item-, location-, or associative-recognition memory judgements made with either high or low confidence. Experiment 3 confirmed that these effects did not depend on whether or not young participants were warned about the subsequent LTM tests before performing the WM task. These dissociations between WM and LTM are inconsistent with the LTM account of latent WM; they are more consistent with the dynamic processing model of WM (*Current Directions in Psychological Science*).

Keywords

Activity-silent; latent; working memory; retrocue; long-term memory; ageing

Received: 19 January 2023; revised: 25 September 2023; accepted: 26 September 2023

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 account? The current research aims to address these questions.

A great deal of related research has focused on the intersection between attention, WM, and LTM (Baddeley, 2000, 2012; Oberauer et al., 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 them using episodic LTM retrieval processes (Atkinson & Shiffrin, 1968; Cowan, 1988, 2001, 2008; D. P. McCabe, 2008; McElree, 2006; Oberauer, 2002; Unsworth & Engle, 2007). Indeed, recent accounts

---

[1]University of Notre Dame, Notre Dame, IN, USA
[2]University of Essex, Colchester, UK

Corresponding author:
Chang Mao Chao, University of Notre Dame, 390 Corbett Family Hall, Notre Dame, IN 46556, USA.
Email: changmac@ucr.edu

suggest that such passively retained “latent” memories are not “in WM” per se—they must be retained in and retrieved “from LTM” (Beukers et al., 2021; Cowan, 2019; Foster et al., 2019; Oberauer & Awh, 2022; Rose, 2020). The current study aimed to distinguish between these different accounts about whether and how episodic LTM retrieval processes are involved in returning information that is retained outside of focal attention during a WM task.

Investigating the roles of attention and LTM 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 prioritise the maintenance of one or more items over other items that were initially encoded and maintained in WM, but, following the retrocue, are to be deprioritised or deleted from WM. Therefore, retrocued items are thought to be held in a higher “state of activation” in the focus of attention than the other deprioritised items, and this typically provides a benefit to WM in terms of memory accuracy, precision, or response times (RTs; for review, see Souza & Oberauer, 2016). For example, Souza et al. (2016) instructed participants to memorise an array of colours followed by either a blank screen (no-cue) or an arrow pointing to the to-be-tested item (retrocue) that was subsequently tested by a probe colour. Their results showed that retrocued items were better recognised than uncued items, which suggests that the retrocues directed participants’ attention to the cued items to prioritise their maintenance and representation in WM.

However, in single retrocue paradigms, the items that are tested on a trial are the ones that are to-be-attended and retained in focal attention throughout the trial. The double-retrocue paradigm is thus particularly useful for characterising the role of LTM processes in WM task performance because it helps to de-confound the role of internally directed *attention* from WM *retention* processes (LaRocque et al., 2015; Lewis-Peacock et al., 2012; Rose et al., 2016). As in the single-retrocue paradigm, the double-retrocue paradigm also presents an initial retrocue that prompts participants to focus on the first-tested item. Thereafter, a second retrocue indicates which item is to be tested next on the trial. Thus, the participant must switch their attention back to the initially uncued/deprioritised/deleted item(s) to reactivate the item(s) that was dropped from focal attention, which is hypothesised to require episodic retrieval processes as in tests of LTM (Cowan, 2008; Unsworth & Engle, 2007).

The double-retrocue paradigm provides important tests of this hypothesis. LaRocque et al. (2015, Experiments 2 and 3) administered a double-retrocue WM task (involving delayed recognition of images of common objects) followed by a surprise, subsequent LTM test of the images from the initial WM task. Assessing subsequent LTM of the items initially processed on the WM task allows one to assess the extent to which episodic LTM retrieval processes were involved in reactivating the deprioritised (passively retained or “latent”) items versus actively retained items on the WM task. This is because practice at using LTM retrieval processes tends to benefit subsequent LTM performance more than processing information held in focal attention or “primary memory” (Loaiza & McCabe, 2012; Mc Cabe, 2008; Rose & Craik, 2012; Rose et al., 2010).

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 attended to and tested first.[1] Participants saw a probe image and responded as to whether it was a match or non-match of the cued image. Then, a second cue and a second probe were presented for the participant to make a match/non-match response about 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. The participants indicated whether or not each image had been presented during the WM task (i.e., old/new item memory). The items were categorised into four conditions based on how they were initially held in different states of prioritisation on the WM task: A–A (attended 1st and attended 2nd), A–U (attended 1st and unattended 2nd), U–A (unattended 1st and attended 2nd), and U–U (unattended 1st and unattended 2nd). Subsequent LTM was assessed with a two-alternative forced-choice recognition test that presented each old item (e.g., mailbox) in its originally studied state (lid closed) and a changed state (lid open) and participants were to recollect and select the one that matched the original state. LTM was compared for these items to assess the extent to which episodic LTM processes were involved in reactivating items held in WM that were dropped from focal attention. Specifically, items from the U–A condition are the ones that were dropped from focal attention and then reactivated later on in the trial. If LTM was engaged in this reactivation process, then subsequent LTM performance should be better for items from the U–A condition than the A–U condition. However, the results showed that subsequent LTM was similar between the U–A and A–U conditions.

Based on this result, the authors concluded that reactivating latent WM items did not involve episodic LTM retrieval processes. This conclusion was consistent with observations from neural recordings and theories based on neurocomputational models (for reviews, see LaRocque et al., 2014; Rose, 2020). However, the LaRocque et al. (2015) study did not address four important issues: (1) the study only assessed performance in healthy young adults (not a population with WM and LTM deficiencies, e.g., older adults), (2) the study did not assess a WM task condition that divided participants' attention during the delay periods to disrupt covert rehearsal processes, (3) the participants did not know that their memory for the items from the initial WM task would be tested later (i.e., encoding for the subsequent LTM test was incidental, not intentional), and (4) the LTM test did not assess memory for different types of details that may accompany retrieval (i.e., memory for associated context bindings, confidence in memory decisions). As explained further on, the current study addresses these potential caveats and related issues.

Many studies, especially those with the double-retrocue paradigm, have shown that active neural representations of uncued items return to baseline during WM delay periods (Lewis-Peacock et al., 2012; Rose et al., 2016), suggesting that the items are not continuously maintained in a sustained, active manner as previously thought (Constantinidis et al., 2018; Fuster & Alexander, 1971; for reviews, see Rose, 2020; Stokes, 2015). 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 an empirically derived baseline-decoding of 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). In brief, the synaptic theory of WM posits that latent items can be represented and briefly retained via short-term plasticity mechanisms in an activity-silent or hidden state. This model also suggests that information in an activity-silent state can be reactivated by nonspecific input to the network. Consistent with this theory, Rose et al. (2016) demonstrated that single pulses of transcranial magnetic stimulation (TMS) could reactivate latent memory 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 et al., 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 (Beukers et al., 2021; Cowan, 2019; Foster et al., 2019; Oberauer & Awh, 2022; Rose, 2020; Rose & Chao, 2022). That is, the neural activity of latent items may drop to baseline because they are no longer retained "in WM" per se; therefore, the latent WM items may be reactivated when these items are subsequently cued and retrieved back into focal attention via episodic LTM retrieval processes. A developing literature suggests that latent WM may be better conceptualised as LTM (Beukers et al., 2021; Cowan, 2019; Foster et al., 2019; Oberauer & Awh, 2022; Rose, 2020; Rose & Chao, 2022). For example, Foster et al. recently proposed an LTM account of latent WM. 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.

As alluded to previously, considerable cognitive research provides some support for this logic. For example, some work suggests that the involvement of LTM processes in WM is greater for items initially studied during complex WM span tasks than simple (short-term memory, STM) span tasks—a phenomenon known as the McCabe effect (e.g., Loaiza & McCabe, 2013; D. P. McCabe, 2008). One explanation of this effect is that the interleaved distraction during a complex WM span task trial displaces the memory items from focal attention and reactivating them requires retrieving them from LTM; in contrast, simple STM span tasks allow participants to continuously 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 et al. (2014, 2015) found behavioural and neural evidence for the involvement of LTM processes in retrieving a single word following just a few seconds of distraction, but the involvement was greater for distraction from a hard versus easy math task. During a final free recall LTM test, the items maintained in the arithmetic conditions were better recalled than other items maintained in the condition without distraction. These results provided evidence for the involvement of LTM processes in WM tasks following hard math, but not following continuous rehearsal (Craik & Watkins, 1973; Rose et al., 2014; for related behavioural research, see Rose & Craik, 2012; Loaiza & Camos, 2016).

These and related findings (Rose & Chao, 2022; Rose et al., 2012, 2015; Slotnick, 2022) collectively suggest that reactivating latent items during WM tasks does involve episodic LTM retrieval processes in many situations. According to a strict view, if information that is initially perceived/encoded and then maintained in the focus of attention "in WM" is not continuously maintained "online" in an active state in focal attention, then returning it to focal attention must involve episodic LTM retrieval, even in the absence of any distraction for just a single item (McElree, 2006).

However, this is at odds with the activity-silent WM account, which posits that latent items may be simply retained in a passive/transient state in WM via short-term—not LTM—processes. A strict view is also at odds with the behavioural, neuroimaging, neurostimulation, and neurocomputational modelling evidence that suggests that such latent items retained outside focal attention are not retrieved with episodic LTM retrieval processes. Episodic LTM retrieval processes and associated neural substrates do not seem to be involved in the short-term retention of information in many situations (Jeneson & Squire, 2012); the involvement has been argued to depend on the type and amount of distraction as well as the type of memoranda and test (Rose & Chao, 2022). That is, the extent to which LTM is involved in WM depends on many contextual variables. Accounting for this variability is what led to the dynamic processing model of WM (Rose, 2020). The goal of the present study is to help distinguish between the activity-silent and LTM accounts of latent WM and thus, the nature of the distinction between WM and LTM, particularly with respect to the role of LTM retrieval in reactivating latent WM.[2]

To summarise, the activity-silent account of latent WM posits that unattended items are still "in WM," whereas the LTM account suggests that the unattended items are retrieved via episodic LTM processes. That is, the LTM account suggests that items are either in focal attention/WM or in LTM; thus, passively retained latent items that are not currently in the focus of attention are held offline in LTM (Beukers et al., 2021; Foster et al., 2019; Rose, 2020). In contrast, neurocognitive models of WM involving embedded processes posit three states of accessibility: (1) active maintenance/representation in focal attention, (2) passive/activity-silent maintenance outside the focus of attention in an intermediate "region of direct access" (Oberauer, 2002, 2009) so that the items are still "in WM" (not "in LTM" per se) readily available for ongoing cognition (Cowan, 1988, 2008, 2019; Jonides et al., 2008; Oberauer, 2002, 2005, 2009), or (3) in LTM which necessitates episodic retrieval. According to these models, item–context bindings may be passively retained and brought back into the focus of attention when they are subsequently cued and reactivated. In contrast, as reviewed previously, The primary goal of the current study is to test these two competing hypotheses by evaluating how the items are maintained and retrieved by comparing WM performance with later LTM of these items from the original WM task. If the LTM account were true, then retrieving items via episodic LTM retrieval processes should leave a "fingerprint" on subsequent LTM performance with detectable differences in LTM for U–A versus A–U items due to the differential involvement of episodic LTM retrieval practice.

The influence of age on the WM/LTM relationship

Another approach to investigating the role of LTM processes in WM has been to examine the nature and sources of adult age differences in WM (Basak & Verhaeghen, 2011; Craik, 1986; Craik & Byrd, 1982; Greene et al., 2020; Morris et al., 1990; Oberauer, 2005; Park, 2000). Older adults' deficits in episodic LTM are large compared with young adults, and other forms of memory (Craik & Rose, 2012), even for those who are ageing normally (Duarte & Dulas, 2020; Nyberg et al., 1996). A fruitful approach has been to examine subsequent LTM of items initially encoded, maintained, and potentially reactivated using episodic LTM processes on an initial WM task (Bartsch et al., 2019; Forsberg et al., 2022; Loaiza & McCabe, 2013; Rose & Craik, 2012; Rose et al., 2010, 2014; Strunk et al., 2019).

WM tasks with attentional prioritisation (e.g., retrocues) help to clarify the effects of maintaining WM items in different states of accessibility. Some studies suggest that older adults are able to use single retrocues as effectively as young adults to guide their attention to the cued item(s) (Gilchrist et al., 2016; Loaiza & Souza, 2018, 2019; Mok et al., 2016; Souza, 2016; Strunk et al., 2019), while others suggest that there are age-related deficits in the use of single retrocues on WM tasks (Duarte et al., 2013; Newsome et al., 2015). However, only one of these studies assessed subsequent LTM for items that were initially maintained on the WM task (Strunk et al., 2019). As reviewed previously, this is critical for assessing how prioritisation in WM might involve episodic LTM processes. Strunk et al. (2019) demonstrated similar retrocue benefits to WM and LTM between young and older adults, but only for item memory—not location memory. Although an informative result, the limited use of single retrocues prevents any assessment of how latent WM items are reactivated. All of the aforementioned studies have examined age differences only in single-retrocue WM tasks, except one (Loaiza & Souza, 2018). Loaiza and Souza (2018) found that healthy older adults showed similar double-retrocue benefits to young adults, thus suggesting that their ability to reactivate latent WM items was preserved. However, without a test of subsequent LTM, it is unclear whether the activity-silent account or LTM account best explains this result. The current study addressed these gaps by investigating whether age deficits to subsequent LTM are specifically observed for items that were reactivated in WM, thus suggesting that episodic LTM processes that are deficient in older age contribute to reactivating latent WM items. Furthermore, these age deficits may be observed most strongly in measures of item–context bindings that are known to be specifically impaired in older age (Naveh-Benjamin et al., 2003; Old & Naveh-Benjamin, 2008).

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

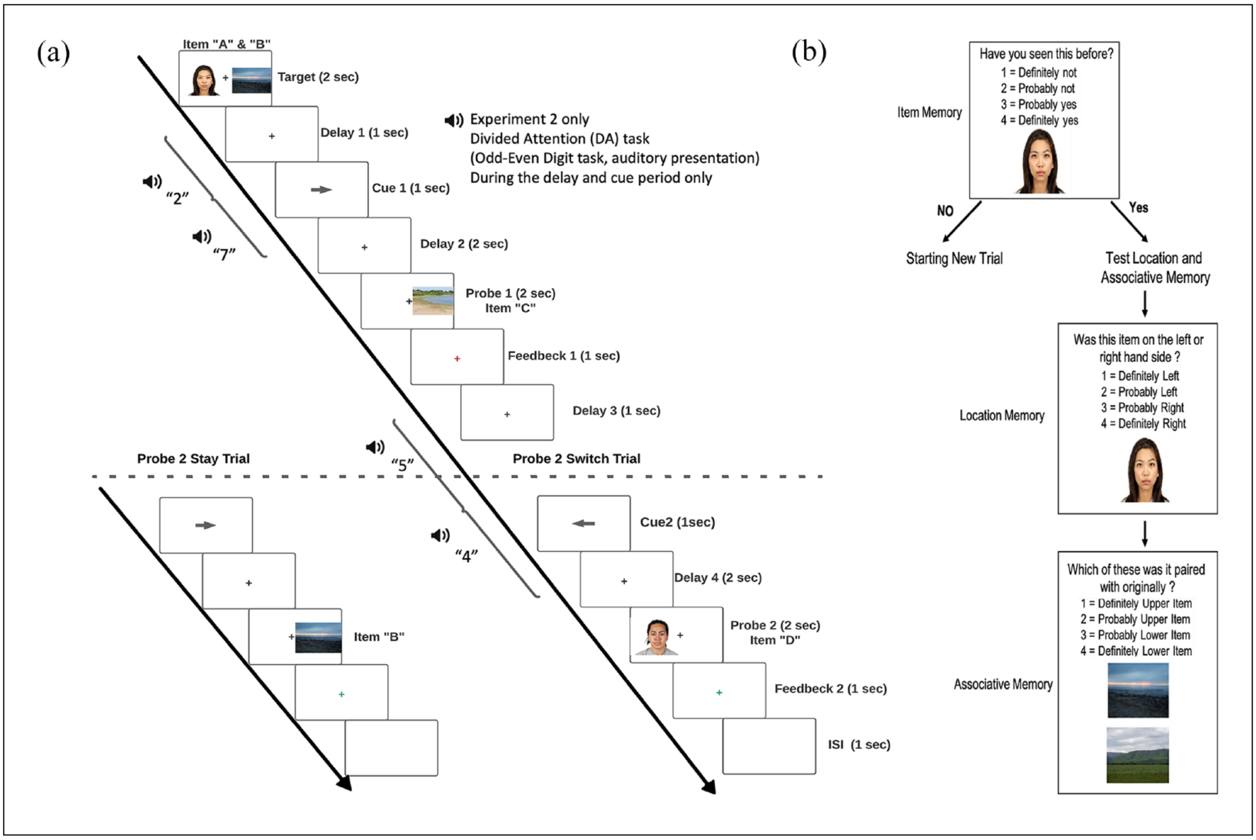

Figure 1. (a) Example of the WM task procedure for either a Probe 2 stay trial (left) or a Probe 2 switch trial (right). For this example of the WM task, if it were a Probe 2 switch trial, the face (Item "A") would be unattended 1st and attended 2nd (UA, and probed by a new item "D") and, thus, the scene (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.

older adults' ability to switch between items in WM (see Figure 1). On each trial, participants encoded two memory items (e.g., a face and place), displayed on the left- and right-hand sides 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. After finishing the double-retrocue WM task, a subsequent LTM test was administered wherein participants made old/new item-recognition judgements followed by judgements about both the associated location (item–location context memory) and the associated item (item–item associative memory) for each of the items judged "old."

Therefore, the current study aimed to replicate and extend aspects of the paradigm and design of LaRocque et al. (2015) (Experiments 2 and 3). The current study extends that research in four ways. First, we examined the effects of dropping and reactivating an item from focal attention during WM maintenance on subsequent LTM in healthy young adults as well as a population with LTM deficiencies (i.e., older adults). Second, we assessed whether these effects to subsequent LTM depended on whether participants were instructed in advance that their LTM would be tested (intentional encoding; Experiments 1 and 2) or not (incidental encoding; Experiment 3). Third, we examined whether engaging young participants in a secondary distractor task (i.e., divided attention [DA]; Experiment 2) during the double-retrocue WM task would moderate the subsequent LTM effects observed under full attention (FA; Experiment 1), assuming that attentional control mechanisms are required to reactivate a retrocued item. Finally, enquiring about the original contextual associations during the subsequent LTM test allowed us to examine whether dropping and reactivating in WM specifically impacts subsequent recollection (i.e., remembering specific details of the associations) versus item recognition, which is thought to be more heavily influenced by the strength of a familiarity signal (Yonelinas, 2001).

Our two key preregistered hypotheses centred on the following logic. If reactivating latent items involves retrieving them using episodic LTM retrieval processes, 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; preregistered hypotheses 1 and 2) both young and older adults' subsequent LTM should differ between items that were initially held in different states of prioritisation on the WM task, particularly between unattended items that were reactivated (U–A condition) versus items that were initially attended but dropped from focal attention later on (A–U condition) during the switch trials. That is, if episodic LTM retrieval 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 (preregistered 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 (preregistered hypothesis 2B).

Following the results of Experiment 1, we recruited additional samples of young adults to participate in Experiments 2 and 3 to determine whether the pattern of results is impacted by DA and incidental encoding, respectively. Specifically, in Experiment 2, we anticipated that young adults whose attention was distracted by a secondary task during the cueing and delay periods of the double-retrocue task should show a selective deficit on prioritising cued over uncued items. Furthermore, Experiment 3 was identical to Experiment 1, except that the young adults were not informed about the subsequent LTM test to ensure that the pattern of results is not attributable to intentional encoding

General methods

Participants

A priori power analyses were conducted using G* Power to estimate the number of participants that would be required 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), which showed a reliable interaction between the effect and age ($n = 27$, power = .95, alpha = .05). Testing at least 54 total participants is required to detect a comparable between-subjects interaction effect size with at least 95% power (assuming a correlation between conditions of .5 and nonsphericity correction at 1). However, because half of the previous studies showed no interaction between retrocueing and age (i.e., Gilchrist et al., 2016; Loaiza & Souza, 2018, 2019; Yi & Friedman, 2014), we aimed to have larger datasets in case any participants' data needed to be excluded for any reason (failure to follow instructions, poor accuracy on the DA task) and to ensure that any failure to detect an effect would not be due to having insufficient power from undersampling. All participants had normal or corrected-to-normal vision and hearing, the ability to discriminate between the colours red and green, and 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 et al., 2010). All participants had a modified-TICS score greater than 34 suggesting that all participants had normal neurocognitive function (Knopman et al., 2010; see Table 1). Performance was significantly higher for the young adults in Experiments 1–3 compared with the older adults, $t(55) = 2.82$, $p = .007$, $t(54) = 2.96$, $p = .004$, and $t(58) = 4.19$, $p < .001$, respectively. Specifically, young adults in all three experiments outperformed older adults on both the initial free recall test, $t(55) = 3.25$, $p = .001$, $t(54) = 3.92$, $p = .0003$ and $t(58) = 5.38$, $p < .001$ respectively, and the final free recall test, $t(55) = 3.15$, $p = .00$, $t(54) = 4.21$, $p = .0001$ and $t(58) = 5.11$, $p < .001$, respectively, on the modified-TICS. These results confirmed that the older adult group had deficits in episodic LTM compared with the young adults.

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

Data exclusion criteria

Videos of the experimental 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 were 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,

Table 1. Sample characteristics. Numbers (participants, in-person/virtual testing, and gender) and mean (age and TICS scores) of all experiments.

	Experiment 1		Experiment 2	Experiment 3
	Young adults, FA	Older adults, FA	Young adults, DA	Young adults, FA/ incidental LTM
No. recruited	30	30	30	35
No. included in analysis	30	27	29	33
No. in-person/virtual testing	10/20	12/18	20/10	35/0
No. of female/male	23/7	16/11	23/6	23/10
Mean age (range; *SD*)	19.87 (18–21; 1.19)	71.92 (64–81; 3.97)	20.10 (18–24; 1.37)	19.26 (18–22; 0.96)
Mean TICS (*SD*)	40.47 (3.22)	38.15 (2.94)	40.48 (2.95)	41.15 (2.60)
No. of responses exceeding task deadline (%)	1.80	3.30	2.10	1.26

TICS: Telephone Interview of Cognitive Status; FA: full attention; DA: divided attention; LTM: long-term memory.

therefore, should be excluded from analyses. For Experiment 1, data for three older adults were excluded from analysis 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 due to the possibility of a tradeoff in performance between the two tasks. One participant was excluded due to performance below 70% on the secondary task. For Experiment 3, data for two participants were excluded from analysis because they did not follow instructions.[3]

Materials and procedure

The experimental tasks were programmed in PsychoPy2 (Peirce et al., 2019) for in-person or online administration via Pavlovia.org. Participants completed the experiment either in-person in our laboratory while abiding by the University- and IRB-approved COVID-19 safety protocols 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 . 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 Supplementary Material).

During the instructions phase, participants were informed that they would perform a "short-term memory" task (Experiments 1–3) followed by a "long-term memory" test of their memory for the items and their original locations and pairings from the initial STM test (only Experiments 1 and 2; see Figure 1). In total, there were 48 trials of the double-retrocue WM task. Each trial started with a central fixation, followed by two items from different categories (e.g., a face and scene) presented to the left (Item A in Figure 1a) and right (Item B) of the fixation cross for 2 s. Thereafter, the items disappeared and a centrally presented arrow pointing to the left or right appeared, indicating which of the two items 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 s to press the "1" or "2" on the computer's keyboard with their left pointer or middle fingers to indicate if the cued item (A or B) matched the probe (C) or not, respectively. The fixation cross turned green or red depending on the accuracy of their response. Then a second arrow appeared pointing to the left or right to indicate which item (the initially presented A or B item) would be tested second. The second cue pointed to either the same item that was just tested (Probe 2 stay) or to the other item (Probe 2 switch) with equal probability. As before, after another short delay, a second probe stimulus (Item D) appeared, to which participants responded regarding whether the probe was a match or non-match of the initially presented stimulus at that location (i.e., the first test probe, Item C, was irrelevant). The number of responses that exceeded the 2 s was small for each group (see Table 1). Therefore, all participants had sufficient time to reach their decision on almost all trials.

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. Immediately after the initial two stimuli (items A and B) were presented, a random series of digits (1–9) was presented auditorily through headphones at a comfortable listening level (~50% of PC volume); participants pressed the "o" key with their right middle finger for odd digits and the "p" key with their right pointer finger for even digits as quickly and as accurately as possible. Participants received feedback immediately via a high (800 Hz) or low (400 Hz) tone for each correct or incorrect response, respectively.

The rate at which the digits were presented was individually determined during a pretesting procedure prior to starting the WM task: the odd–even digit task was presented on its own, and participants indicated if the digit was odd or even as quickly and accurately as possible. 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 s from the digit onset. 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 RT 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 RT of the last three reversals from the pretest and adding two standard deviations. We first confirmed that young adults in Experiment 2 performed this DA task at a high level before proceeding with the WM task trials. Average accuracy on the secondary task during the WM task trials was 81.03% correct ($SD=6.4\%$). Given this individual titration procedure and the fixed 4 s between the target and probe, it was possible for participants to respond to either two ($N=28$) or three digits ($N=2$) during each cue and delay period. However, control analyses confirmed that the pattern of results did not differ when excluding participants with three digits (see Supplementary Material). All other details were the same as Experiment 1.

Following the WM task for all three experiments, participants were administered the TICS, lasting approximately 5 min. The participants then completed the subsequent LTM test, comprising "old" items from the initial 48 trials of the WM task, and an equal number of novel, lure items from each category (96 trials in total) presented one at a time in a random order. The old items evenly represented the four different conditions of the WM task 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.[4]

For each image, the participants indicated whether they thought the item was "old" (a to-be-remembered item from the WM task) or "new" (not presented during the WM task) on a four-point confidence scale (definitely old, probably old, probably new, or definitely new). Participants had as much time as they needed to make each response. Responses were recorded as hits (old items called definitely or probably old) and false alarms (new items called definitely or probably old), and item memory accuracy was scored as hits—false alarms. For all "old" judgements, participants used a similar four-point confidence scale to first indicate the side of the screen on which the item was initially presented to measure their location memory, followed by indicating which of two images (respectively, presented in the upper- and lower-half of the screen) was originally presented with the item in question during the WM task to measure their associative memory. 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.

For the stimuli (pictures of faces, scenes, and names), face stimuli were obtained from the Chicago Face Database (Singh et al., 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 et al., 2017). Scenes that would not be readily identifiable or recognised by our participants were selected. The NAfME 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 America-dwelling, English-speaking participants.

Data analysis and predictions

With regard to the WM task, accuracy was assessed as a function of age group (between-subjects factor) and probe-type condition (Probe 1, Probe 2 stay, and Probe 2 switch; within-subjects factor) in a mixed analysis of variance (ANOVA). We also used JASP to conduct Bayesian analysis of variance (BANOVA; Rouder et al., 2012) and Bayesian *t* test (Rouder et al., 2009) to calculate Bayes factors (BF) and estimate the strength of evidence favouring the null (BF_{01}) versus alternative (BF_{10}) hypothesis (van Doorn et al., 2021).

RT data and analyses are reported in the supplementary 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 Probes 1 and 2 stay trials than Probe 2 switch trials, because the former 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 versus 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. Note that it was possible for the same non-match probe stimulus to be presented on both Probes 1 and 2 in a subset of stay trials. Investigation of these repeated lure 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 preregistered hypotheses and Xu et al., in press). For the present purposes, these trial types were not included in the analyses reported further on.

Furthermore, we predicted that overall lower WM accuracy in older than young adults could stem from two possibilities: first, if both items are in the focus of attention under full attention, then an age effect may indicate a deficit in switching attention between them, and retrieval from LTM may not be involved. Conversely, if reactivating latent items requires retrieval from LTM, then older adults with LTM deficits should perform worse than young adults under FA, particularly on Probe 2 switch trials, and on the subsequent LTM tests. We expected that there would be an interaction between age and probe-type, such that the age difference on Probes 1 and 2 stay trials would be smaller than the age difference on Probe 2 switch trials. Note that this prediction is inconsistently found in the literature; the source of this variability is the focus of our related article (Xu et al., in press).

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 non-existent 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 different age-groups using a mixed ANOVA for item memory, location memory, and associative memory (collapsed over "definitely" and "probably" judgements). Follow-up *t*-tests were used to test the a priori hypotheses described previously. Our preregistered hypotheses did not specify any a priori predictions about differences in the effects of reactivation, age, or DA on subsequent LTM as a function of confidence. However, we conducted these analyses on item, location, and associative memory separated by high- and low-confidence judgements to supplement the analyses of memory overall because high-confidence judgements, which should capture recollection (M. K. Johnson et al., 1988), may be more sensitive to potential UA–AU differences than low confidence judgements, which capture familiarity-based responding. The analyses on LTM were also repeated to examine performance separately for items recognised with high ("definitely") versus low ("probably") confidence (see supplemental results).

Note that the most appropriate contrast for subsequent LTM was between UA and AU items that were correctly retrieved initially because UA items were the items that were to be dropped from focal attention and potentially retrieved with LTM processes, whereas AU items were the items that were initially retained in and retrieved directly from focal attention but dropped from focal attention later on, so the amount of time the UA and AU items spent in focal attention was matched. Therefore, any performance differences between the UA and AU items on subsequent LTM tests could be attributed to differences in the nature of the retrieval processes involved in WM. The AA items, with double retrieval practice, and UU items, with no retrieval practice, provide control conditions that help with understanding the nature of WM retrieval and its impacts on subsequent LTM. Better LTM for UA than AU items would suggest that reactivating UA items involved more episodic LTM retrieval practice than AU items. Also note that, as in previous research (e.g., Rose & Craik, 2012; Rose et al., 2014), we analysed those items that were initially retrieved correctly on the WM task to ensure that any differences in LTM performance were not due to differences in initial WM accuracy (for a similar approach, see Forsberg et al., 2022).

To contrast the effects of reactivating latent items, age, and DA on WM and subsequent LTM, effect sizes (partial eta squared) of the main effects were calculated to facilitate comparison across the different tests. The reactivation effect on WM was represented by the main effect of switching (Probe 2 switch vs. Probe 2 stay) on WM with a two-way mixed ANOVA for the groups (Young FA, Old FA, Young DA [Experiment 2]). The reactivation effect on subsequent LTM was represented by the main effect of critical comparison between UA versus AU items on each LTM test with a two-way mixed ANOVA for the groups (Young FA, Old FA, Young DA [Experiment 2]). The ageing effect on WM was represented by the main effect of age-group (Young FA, Old FA) on overall WM performance (Probe 1, Probe 2 stay, Probe 2 switch) with a two-way mixed ANOVA. The ageing effect on LTM was represented by the main effect of age-group (Young FA, Old FA) on overall LTM performance (AA, AU, UA, UU) for each test with separate two-way mixed ANOVAs. These age effects were compared with the age effect on a separate LTM test with a two-way mixed ANOVA on final free recall performance on the TICS neuropsychological screening test with an independent samples *t*-test and converting the effect size (Cohen's *d*) to partial eta squared by dividing Cohen's *d*^2 × N by Cohen's *d*^2 × N + N − 1 github source (2020): https://haiyangjin.github.io/2020/05/eta2d/. The DA effect on WM was represented by the main effect of group (Young FA, Young DA) on overall WM performance (Probe 1, Probe 2 stay, Probe 2 switch) with a two-way mixed ANOVA. The DA effect on subsequent LTM was represented by the main effect of group (Young FA, Young DA) on overall LTM performance (AA, AU,

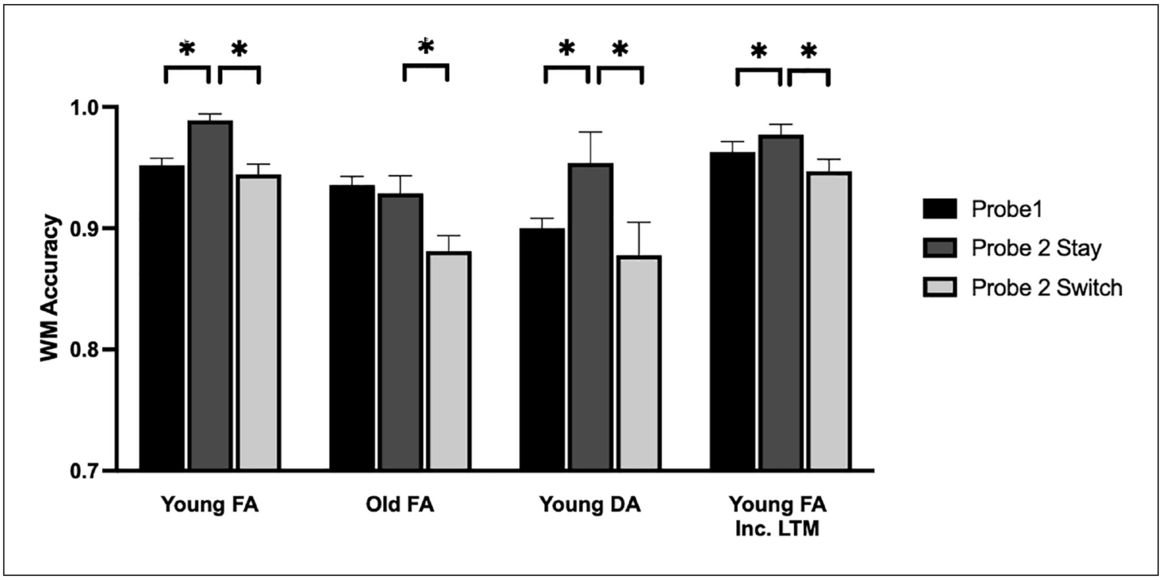

Figure 2. Average accuracy on the WM task for the three types of trial probes in Experiment 1 (Young FA and Old FA), Experiment 2 (Young DA) and Experiment 3 (Young FA Inc. LTM).
FA: full attention; DA: divided attention; Inc; LTM: incidental encoding for subsequent LTM test.
Error bars indicate the standard error of mean.
$^*p < .05$ from two-tailed, paired-sample *t*-tests.

UA, UU) for each LTM test with a two-way mixed ANOVA.

Data availability statement

This study's design and hypotheses were preregistered; see https://osf.io/z9cgq/. All data and study materials have been made publicly available and can be accessed at Open Science Framework (https://osf.io/ztqx8/).

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 contextualise the interpretation of the LTM data.

WM accuracy results

Average performance on the WM task for all three experiments is presented in Figure 2 and the omnibus ANOVAs are presented in Table 2.

Experiment 1. To test our pre-registered hypothesis 1, we compared age differences of the items that were possibly reactivated via LTM retrieval processes (Probe 2 switch trials, UA items) to that of the items that were presumably maintained in focal attention throughout the trial (Probe 2 stay trials, AA items). If reactivating items on Probe 2 switch trials involves retrieval from LTM, which is deficient in older age, then the age difference should be larger on Probe 2 switch trials compared with Probe 2 stay trials that presumably do not require retrieval from LTM. Against this hypothesis, however, the size of the age difference on the Probe 2 switch trials (Young FA *M*—Old FA $M = 0.06$; 95% CI = [0.018 to 0.109]) was similar to that of Probe 2 stay trials (Young FA *M*—Old FA $M = 0.06$; 95% CI = [0.03 to 0.09]).

Thus, older adults performed significantly worse overall compared with young adults, but the cost of switching versus staying was the same for both young, $t(29) = 4.65$, $p < .001$, $BF_{10} = 368.55$, and older adults, $t(26) = 3.29$, $p = .003$, $BF_{10} = 13.57$, suggesting that young and older adults with FA were similarly able to use the retrocues to prioritise the cued item over the uncued item. A significant interaction emerged because young adults showed improved WM accuracy on Probe 2 stay trials relative to Probe 1 trials, $t(29) = 6.55$, $p < .001$, $BF_{10} = 45{,}499.43$, perhaps due to being tested on the same item twice and receiving feedback following Probe 1; in contrast, older adults did not benefit from repeated testing with feedback, $t(26) = 0.56$, $p = .58$, $BF_{01} = 4.26$. Instead, older adults performed worse on Probe 2 switch trials compared with Probe 1, $t(26) = 3.57$, $p = .001$, $BF_{10} = 25.20$, whereas there was no difference between Probe 1 and Probe 2 switch accuracy for young adults with FA, $t(29) = 0.66$, $p = .51$, $BF_{01} = 4.20$.

Experiment 2. We further investigated whether a secondary DA task during the maintenance interval would impair young adults' ability to prioritise cued items relative to older adults. As evident in Figure 2 and Table 2, a

Table 2. Results of ANOVAs and BANOVAs of main effects (Group and Condition) and interactions (Group × Condition) for all experiments on WM accuracy.

Measure	Exp.	Analysis	Main effect of group	Main effect of condition	Group × Condition interaction
WM accuracy	1	2 (Group: Young FA, Old FA) × 3 (Condition: Probe 1, Probe 2 stay, Probe 2 switch)	$F(1,55)=9.83$, $p=.003$, $\eta p^2=.15$, $BF_{10}=13.78$	$F(2,110)=15.90$, $p<.001$, $\eta p^2=.22$, $BF_{10}=7{,}001$	$F(2,110)=4.97$, $p=.012$, $\eta p^2=.08$, $BF_{10}=4.90$
	2	2 (Group: Young DA, Old FA) × 3 (Condition: Probe 1, Probe 2 stay, Probe 2 switch)	$F(1,54)=0.07$, $p=.78$, $\eta p^2=.00$, $BF_{01}=3.73$	$F(2,108)=17.42$, $p<.001$, $\eta p^2=.24$, $BF_{10}=35{,}202$	$F(2,108)=3.84$, $p=.03$, $\eta p^2=.07$, $BF_{10}=2.16$
		2 (Group: Young FA, Young DA) × 3 (Condition: Probe 1, Probe 2 stay, Probe 2 switch)	$F(1,57)=14.70$, $p<.001$, $\eta p^2=.21$, $BF_{10}=76.28$	$F(2,114)=24.43$, $p<.001$, $\eta p^2=.29$, $BF_{10}=2{,}673{,}930$	$F(2,114)=1.75$, $p=.19$, $\eta p^2=.03$, $BF_{01}=3.27$
	3	2 (Group: Young Intentional, Young Incidental) × 3 (Condition: Probe 1, Probe 2 stay, Probe 2 switch)	$F(1,61)<0.01$, $p=.96$, $\eta p^2<.001$, $BF_{01}=3.68$	$F(2,122)=19.74$, $p<.001$, $\eta p^2=.245$, $BF_{10}=21{,}819.95$	$F(2,122)=1.72$, $p=.184$, $\eta p^2=.027$, $BF_{01}=2.81$

ANOVA: analysis of variance; BANOVA: Bayesian analysis of variance; WM: working memory; FA: full attention; BF: Bayes Factors; DA: divided attention.

significant interaction emerged, indicating that the group difference (old FA vs. young DA) was significant for Probe 1 trials, $t(54)=2.26$, $p=.028$, $BF_{10}=2.141$, but not for Probe 2 stay trials, $t(54)=1.49$, $p=.142$, $BF_{01}=1.477$ or Probe 2 switch trials, $t(54)=0.12$, $p=.905$, $BF_{01}=3.68$. The size of the group difference on Probe 2 switch trials (Old FA mean—Young DA $M=0.003$; 95% CI = [−0.05 to 0.06]) was similar to that of Probe 1 trials (Old FA mean—Young DA $M=0.035$; 95% CI = [−0.004 to 0.07]) and Probe 2 stay trials (Old FA mean—Young DA $M=-0.02$; 95% CI = [−0.06 to 0.008]). For completeness, exploratory analyses comparing the young DA group and young FA group revealed that the young DA group had higher accuracy on Probe 2 stay trials than both Probe 1, $t(28)=4.85$, $p<.001$, $BF_{10}=573.49$ and Probe 2 switch trials, $t(28)=4.37$, $p<.001$, $BF_{10}=175.57$. Thus, although they performed more poorly overall than young adults with FA, young adults with DA were still able to similarly use both the retrocues to prioritise the cued items and the feedback following Probe 1 to benefit their performance on Probe 2 stay trials.

Experiment 3. Adding the young adult group with FA and incidental LTM encoding allowed us to assess the effect of incidental encoding on WM and subsequent LTM performance. The results showed that WM performance was similar regardless of being informed about the subsequent LTM test (intentional LTM encoding, Exp. 1) or not (incidental LTM encoding, Exp. 3). As in Experiment 1, the young FA group with incidental LTM encoding in Experiment 3 showed higher WM accuracy on Probe 2 stay trials than Probe 1, $t(32)=7.95$, $p<.001$, $BF_{10}=272{,}000$ and Probe 2 switch trials, $t(32)=8.24$, $p<.001$, $BF_{10}=577{,}200$. WM performance did not differ between Probes 1 and 2 switch trials, $t(32)=1.67$, $p=.10$, $BF_{01}=1.53$.

Subsequent LTM results

We first analysed the average performance on the subsequent item, location, and associative recognition tests for correct items from the initial WM task in the different conditions (AA, AU, UA, UU) for the young and older adults with FA (Experiment 1), young adults with DA (Experiment 2), and young adults with FA and incidental LTM encoding (Experiment 3; see Supplementary Table 1). For some control analyses, we also repeated the LTM analyses for all items, irrespective of whether or not they were correctly retrieved on the initial WM task (see Supplementary Table 2). Then we repeated the analyses on the data separated by high- and low-confidence responses (see Table 4).

Experiment 1. To test our preregistered hypothesis 2A, we compared the mean difference between items that were dropped from focal attention and reactivated (UA) and items that were retained in focal attention and then dropped from maintenance (AU) on the hit—false alarm rate for the item recognition test, and the (two-alternative forced choice) location and associative recognition tests for those items recognised as "old." As evident in Table 3, there were no significant differences between the UA and AU conditions for item (−0.001, 95% CI = [−0.07 to 0.07]), location (0.03, 95% CI = [−0.09 to 0.03]), or associative

Table 3. Subsequent LTM results of ANOVAs and BANOVAs of main effects (Group and Condition) and interactions (Group × Condition) for all (both high and low confidence) responses on the item (hits-false alarms), location and associative two-alternative forced choice tests (2AFC) recognition memory tests for all three experiments.

Measure	Exp.	Analysis	Main effect of group	Main effect of condition	Group × Condition interaction
LTM: item memory	1	2 (Group: Young FA, Old FA) × 4 (Condition: AA, AU, UA, UU)	$F(1,55)=0.095$, $p=.759$, $\eta p^2=.002$, $BF_{01}=3.21$	$F(3,165)=2.57$, $p=.056$, $\eta p^2=.045$, $BF_{01}=2.82$	$F(3,165)=1.31$, $p=.266$, $\eta p^2=.023$, $BF_{01}=4.86$
	2	2 (Group: Young DA, Old FA) × 4 (Condition: AA, AU, UA, UU)	$F(1,54)=0.001$, $p=.97$, $\eta p^2<.001$, $BF_{01}=3.36$	$F(3,162)=3.39$, $p=.019$, $\eta p^2=.059$, $BF_{10}=1.62$	$F(3,162)=0.58$, $p=.63$, $\eta p^2=.011$, $BF_{01}=10.95$
		2 (Group: Young FA, Young DA) × 4 (Condition: AA, AU, UA, UU)	$F(1,57)=0.096$, $p=.76$, $\eta p^2=.002$, $BF_{01}=3.52$	$F(3,171)=2.60$, $p=.054$, $\eta p^2=.044$, $BF_{01}=2.12$	$F(3,171)=2.43$, $p=.07$, $\eta p^2=.041$, $BF_{01}=1.41$
	3	2 (Group: Young Intentional, Young Incidental) × 4 (Condition: AA, AU, UA, UU)	$F(1,61)=0.84$, $p=.36$, $\eta p^2=.014$, $BF_{01}=2.65$	$F(3,183)=8.96$, $p<.001$, $\eta p^2=.128$, $BF_{01}=1665.9$	$F(3,183)=1.30$, $p=.28$, $\eta p^2=.032$, $BF_{01}=4.00$
LTM: location memory	1	2 (Group: Young FA, Old FA) × 4 (Condition: AA, AU, UA, UU)	$F(1,55)=1.19$, $p=.28$, $\eta p^2=.021$, $BF_{01}=2.93$	$F(3,165)=0.62$, $p=.604$, $\eta p^2=.011$, $BF_{01}=21.71$	$F(3,165)=1.04$, $p=.375$, $\eta p^2=.019$, $BF_{01}=6.09$
	2	2 (Group: Young DA, Old FA) × 4 (Condition: AA, AU, UA, UU)	$F(1,54)=0.72$, $p=.40$, $\eta p^2=.013$, $BF_{01}=4.14$	$F(3,162)=0.78$, $p=.51$, $\eta p^2=.014$, $BF_{01}=16.31$	$F(3,162)=0.50$, $p=.69$, $\eta p^2=.009$, $BF_{01}=11.39$
		2 (Group: Young FA, Young DA) × 4 (Condition: AA, AU, UA, UU)	$F(1,57)=4.00$, $p=.050$, $\eta p^2=.065$, $BF_{01}=1.20$	$F(3,171)=0.28$, $p=.84$, $\eta p^2=.005$, $BF_{01}=3.23$	$F(3,171)=0.69$, $p=.56$, $\eta p^2=.012$, $BF_{01}=9.06$
	3	2 (Group: Young Intentional, Young Incidental) × 4 (Condition: AA, AU, UA, UU)	$F(1,61)<0.01$, $p=.95$, $\eta p^2<.001$, $BF_{01}=5.16$	$F(3,183)=0.76$, $p=.52$, $\eta p^2=.012$, $BF_{01}=18{,}59$	$F(3,183)=0.76$, $p=.52$, $\eta p^2=.01$, $BF_{01}=18.30$
LTM: associative memory	1	2 (Group: Young FA, Old FA) × 4 (Condition: AA, AU, UA, UU)	$F(1,55)=0.39$, $p=.54$, $\eta p^2=.007$, $BF_{01}=4.80$	$F(3,165)=2.70$, $p=.048$, $\eta p^2=.047$, $BF_{10}=0.88$	$F(3,165)=0.66$, $p=.59$, $\eta p^2=.012$, $BF_{01}=9.82$
	2	2 (Group: Young DA, Old FA) × 4 (Condition: AA, AU, UA, UU)	$F(1,54)=0.03$, $p=.87$, $\eta p^2=.016$, $BF_{01}=4.92$	$F(3,162)=1.15$, $p=.33$, $\eta p^2=.021$, $BF_{01}=10.09$	$F(3,162)=1.24$, $p=.30$, $\eta p^2=.022$, $BF_{01}=4.91$
		2 (Group: Young FA, Young DA) × 4 (Condition: AA, AU, UA, UU)	$F(1,57)=0.56$, $p=.46$, $\eta p^2=.010$, $BF_{01}=4.17$	$F(3,171)=2.10$, $p=.10$, $\eta p^2=.035$, $BF_{01}=3.24$	$F(3,171)=2.17$, $p=.09$, $\eta p^2=.037$, $BF_{01}=1.66$
	3	2 (Group: Young Intentional, Young Incidental) × 4 (Condition: AA, AU, UA, UU)	$F(1,61)=3.04$, $p=.09$, $\eta p^2=.04$, $BF_{01}=1.47$	$F(3,183)=0.19$, $p=.66$, $\eta p^2=.006$, $BF_{01}=0.18$	$F(3,183)=1.68$, $p=.17$, $\eta p^2=.03$, $BF_{01}=3.08$

LTM: long-term memory; ANOVA: analysis of variance; BANOVA: Bayesian analysis of variance; FA: full attention; 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; BF: Bayes factors; DA: divided attention.

memory (−0.04, 95% CI = [−0.02 to 0.10]). Thus, subsequent LTM was not better for UA than AU items overall.

To test our preregistered 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, *M*—Young FA, *M*=−0.03, 95% CI = [−0.08 to 0.13]; AU: Old FA, *M*—Young FA, *M*=−0.03, 95% CI = [−0.12 to

Table 4. Mean proportion correct and standard error of the mean (SEM) on the subsequent LTM tests for high-confidence responses on the item (hits-false alarms), location, and associative two-alternative forced choice tests (2AFC) recognition memory tests for items that were held in attended (A) or unattended (U) states during the first and second delay periods of the initial double-retrocue WM task for the Young FA and Old FA (Experiment 1), Young DA (Experiment 2) and Young FA Incidental (Inc.) LTM encoding (Experiment 3) groups. The critical preregistered comparison of UA–AU items (the reactivation effect) includes 95% CI.

Mean (SEM)	Item memory					Location memory					Associative memory				
	Hits—FA					2AFC					2AFC				
	AA	AU	UA	UU	UA–AU (95% CI)	AA	AU	UA	UU	UA–AU (95% CI)	AA	AU	UA	UU	UA–AU (95% CI)
Young FA	.60(.04)	.56(.04)	.60(.04)	.53(.06)	.04 [–.08, .12]	.82(.04)	.81(.05)	.84(.04)	.76(.05)	.03 [–.18, .09]	.92(.03)	.89(.03)	.89(.03)	.82(.04)	.00 [–.07, .11]
Old FA	.51(.06)	.51(.04)	.50(.05)	.45(.04)	–.01 [–.05, .09)	.65(.06)	.69(.05)	.77(.05)	.72(.05)	.08 [–.22, .16)	.83(.04)	.90(.03)	.86(.04)	.76(.05)	–.04 [–.06, .15]
Young DA	.54(.06)	.62(.05)	.49(.05)	.45(.05)	–.13 [–.27, –.01)	.74(.06)	.74(.06)	.77(.06)	.68(.07)	.03 [–.31, .07]	.89(.03)	.84(.04)	.88(.03)	.89(.03)	.04 [–.04, .15]
Young FA Inc.	.64(.06)	.64(.07)	.60(.06)	.51(.06)	–.04 [–.15, .03]	.92(.03)	.88(.04)	.89(.03)	.86(.04)	.01 [–.09, .13]	.81(.03)	.67(.04)	.80(.03)	.67(.04)	.13 [–.02, .25]

LTM: long-term memory; 2AFC: two-alternative forced choice; WM: working memory; FA: full attention; DA: divided attention; UA: the tested item was unattended 1st and attended 2nd; AU: the tested item was attended 1st and unattended 2nd; CI: confidence interval; SEM: standard error of the mean; AA: the tested item was attended 1st and attended 2nd; UU: the tested item was unattended 1st and unattended 2nd.

0.06]), location memory test (UA: Old FA, M—Young FA, $M=-0.05$; 95% CI = [−0.13 to 0.05]; AU: Old FA, M—Young FA, $M=0.001$, 95% CI = [−0.09 to 0.09]), and associative memory test (UA: Old FA, M—Young FA, $M=-0.03$, 95% CI = [−0.05 to 0.13]; AU: Old FA, M—Young FA, $M=-0.06$, 95% CI = [−0.07 to 0.11]).[5] As evident in Table 3, the lack of main effect of condition, age group, and the interaction between them suggested that the way the items were initially maintained and retrieved on the WM task did not influence subsequent item recognition memory and location memory. Item memory and location memory also did not differ between young and older adults. For associative memory, the main effect of age group and interaction between age group and condition were not significant. The main effect of condition was significant because memory was better for AA items that were tested twice compared with UU items that were never tested.

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. However, as evident in Table 3, the lack of group effects and the interactions between group and condition suggested that the young adults with DA performed similarly to both the young and older adults with FA. Although dividing attention during WM maintenance and cueing impaired WM performance overall, 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. The implications of these striking dissociations between latent WM and LTM are discussed further on.

Experiment 3. Finally, adding an additional group of young adults who were not informed about the LTM test allowed us to assess the impact of intentional versus incidental encoding on the pattern of LTM results. As evident in Table 3, the lack of any main effect of group and interaction between group and condition for the item, location and associative memory tests suggested that whether or not participants were forewarned about the subsequent LTM tests did not affect participants' performance overall.

Analysis of high- and low-confidence judgements. We compared the reactivation effect—the mean difference between UA and AU items—for high-confidence responses on the item, location, and associative recognition tests. As may be seen in Table 4, the subsequent LTM was not better for the reactivated (UA) item compared with the control (AU) item for any group or test, even without correcting for multiple comparisons.

Next, we compared the mean age difference between the older adult and young adult groups on the item, location, and associative recognition tests. As evident in Table 4, the mean age difference was similar between the critical UA and AU items on the item memory, location memory, and associative memory test. For item memory, older adults' performance was not worse than either the young FA group's: $F(1,55)=1.98$, $p=.165$, $\eta p^2=.035$, $\mathrm{BF}_{01}=1.60$, or young DA group's: $F(1,54)=2.22$, $p=.142$, $\eta p^2=.04$, $\mathrm{BF}_{01}=2.69$. There were no significant interactions; old versus young FA: $F(3,165)=0.82$, $p=.486$, $\eta p^2=.015$, $\mathrm{BF}_{01}=7.01$; old versus young DA: $F(3,162)=1.07$, $p=.365$, $\eta p^2=.019$, $\mathrm{BF}_{01}=6.00$.

For location memory, older adults were significantly worse than the young FA group: $F(1,42)=3.88$, $p=.046$, $\eta p^2=.084$, $\mathrm{BF}_{10}=2.77$, but not the young DA group: $F(1,57)=0.30$, $p=.584$, $\eta p^2=.005$, $\mathrm{BF}_{01}=3.86$. There were no interactions between group and condition; old versus young FA: $F(3,126)=1.07$, $p=.367$, $\eta p^2=.025$, $\mathrm{BF}_{01}=5.63$; old versus young DA: $F(3,171)=1.22$, $p=.303$, $\eta p^2=.021$, $\mathrm{BF}_{01}=5.29$. Thus, for high-confidence location memory decisions, there was the expected age deficit between older adults and young adults with FA.

For associative memory, there were no main effects of group: old versus young FA: $F(1,43)=1.80$, $p=.186$, $\eta p^2=.04$, $\mathrm{BF}_{01}=1.75$; old versus young DA: $F(1,38)=1.03$, $p=.317$, $\eta p^2=.026$, $\mathrm{BF}_{01}=3.03$; or interactions: old versus young FA: $F(3,129)=0.89$, $p=.449$, $\eta p^2=.022$, $\mathrm{BF}_{01}=6.37$; old versus young DA: $F(3,114)=2.31$, $p=.08$, $\eta p^2=.057$, $\mathrm{BF}_{01}=1.01$. As discussed further on, we suspect that this unexpected lack of an age difference in associative LTM was due to the fact that the old item that was paired with the target stimulus on the WM task was presented alongside a new item on the two-alternative forced choice associative memory test, so the correct item could be selected based on the strength of the familiarity signal.

For Experiment 3, this control experiment compared a young FA group (with incidental LTM encoding) with the young FA group from Experiment 1 (with intentional LTM encoding). There was no main effect of group or a group by condition interaction for item memory, group effect: group × condition interaction: $F(3,183)=0.23$, $p=.88$, $\eta_p^2=.004$, $\mathrm{BF}_{01}=17.75$ or location memory, group effect: $F(1,45)=0.24$, $p=.62$, $\eta_p^2=.007$, $\mathrm{BF}_{01}=1.11$, condition $F(3,135)=0.54$, $p=.66$, $\eta_p^2=.07$, $\mathrm{BF}_{01}=18.86$; group × condition interaction: $F(3,145)=0.11$, $p=.93$, $\eta_p^2=.002$, $\mathrm{BF}_{01}=16.09$). However, for associative memory, there was a main effect of group, $F(1,56)=14.25$, $p<.001$, $\eta_p^2=.20$, $\mathrm{BF}_{10}=65.71$; the group by condition interaction was not significant, $F(3,168)=1.20$, $p=.31$, $\eta_p^2=.02$, $\mathrm{BF}_{01}=6.00$. The main effect of group suggests that intentional LTM encoding did result in better high-confidence associative memory performance overall compared with incidental LTM encoding.

To summarise the LTM results as a function of confidence, there were no group by condition interactions that differed for high- versus low-confidence judgements. There were no group differences (except for old vs. young

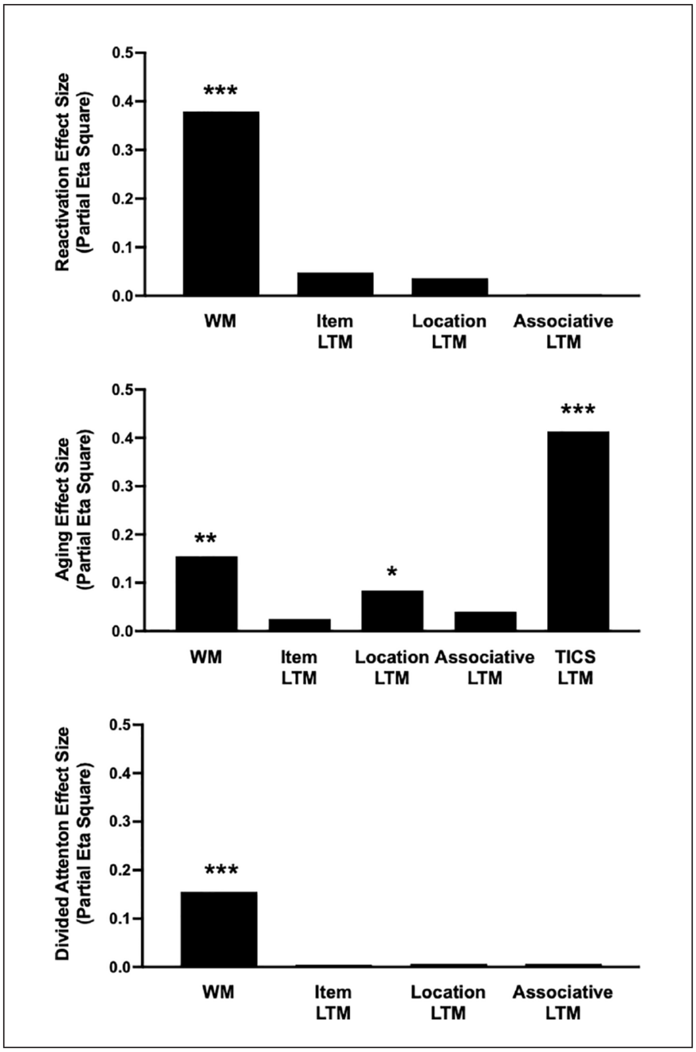

Figure 3. Effect sizes for the three manipulated factors on the double-retrocue working memory (WM) task and on the subsequent item-, location-, and associative recognition long-term memory (LTM) tests for high-confidence responses to correctly retrieved items on the initial WM task: The reactivation effect (top: AU vs. UA in LTM; AA vs. UA in WM), the age-group effect (middle: Experiment 1 Young FA vs. Old FA; including delayed recall on the TICS neuropsychological screening test), and the divided attention effect (bottom: Experiment 1 Young FA vs. Experiment 2 Young DA).
*$p < .05$. **$p < .01$. ***$p < .001$.

FA on location memory) and there were no age differences in mean hit-FA rates between the critical AU or UA items for high- and low-confidence judgements on the item memory test. Therefore, the main conclusions regarding LTM in this article largely do not need to be qualified by these analyses separated by the level of confidence.

Comparison of effect sizes on WM and LTM. To summarise the results and facilitate comparison of the size of the effects of reactivating latent items, age, and DA on WM and LTM performance, the effect sizes (partial eta squared) for the main effects of each of these factors were derived from separate mixed ANOVAs on the WM and high-confidence LTM data for each test. Note that to fairly compare LTM performance for items from conditions that differed in performance on the WM task, the LTM data are for items that were correctly retrieved on the initial WM task. These effect sizes are shown in Figure 3. The reactivation effect size on WM was large whereas the effect sizes on subsequent high-confidence item-, location-, and associative-LTM were all small.[6] Similarly, the age effect size was large on WM, but for subsequent LTM of items correctly retrieved on the initial WM task, the main effect of age-group was only significant for location memory. To contrast the relatively small effect of age on subsequent LTM of correctly retrieved WM items, the age effect on the final free recall test of the TICS neuropsychological screening test is also shown in Figure 3. This confirms that this particular sample of older adult participants did have the expected deficit in episodic LTM on a standardised neuropsychological screening test. As discussed further on, whereas the lack of an age effect on item memory was not surprising, it was for associative memory. This was likely due to the lures on the 2AFC test being novel; had the test presented the paired item and an old item from a different pair, the age deficit in associative memory would likely have been larger. For the DA effect (bottom), again the effect on WM was large, whereas the effect on item-, location-, and associative-memory was small. To sum up, although reactivation, age, and DA effects all showed significant effects on initial WM performance, the effects generally had little impact on subsequent LTM performance. As discussed further on, these dissociations are inconsistent with the LTM account of latent WM.

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 to be retained and reactivated outside of focal attention, (2) older adults versus young adults, and (3) young adults with divided versus full attention. That WM performance was worse for Probe 2 switch than stay trials suggested that these items were retained outside of focal attention by both young and older adults and may have required LTM retrieval processes to reactivate them. However, even though switching and reactivating the latent items, age, and DA all had substantial effects on initial WM performance, these factors had little systematic effects on subsequent LTM performance, regardless of whether LTM was measured by item memory, location memory, or associative memory, with either high confidence (see Table 1), or overall (see Supplementary Table 2).[7] These dissociations between WM and LTM performance suggest that retaining and reactivating latent items on this WM task did not strongly implicate episodic LTM retrieval processes. We discuss this interpretation and its implications for WM theory further on.

No impact of switching attention in WM on subsequent 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; D. P. McCabe, 2008). The results from this double-retrocue WM task showed that subsequent LTM was generally not affected by the way that participants maintained and reactivated the items on the initial WM task. This suggests that the latent items may not have been retained and reactivated with episodic LTM 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 (with similar presentation and retention intervals) while recording and decoding patterns of functional magnetic resonance imaging (fMRI) and electroencephalogram (EEG) activity associated with the 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 signalled 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 both behaviour (accuracy and response time) and 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 (Beukers et al., 2021; Foster et al., 2019; Rose, 2020). This interpretation is consistent with some embedded process models of WM (Cowan, 1995) and buffer-based models of WM (Camos & Barrouillet, 2011).

Indeed, Foster et al. (2019) recently suggested that a lack of delay period neural activation of an item during a WM task indicates that the item is not "in WM," but is "in LTM" instead. However, the lack of differences between unattended and attended items on all of the LTM tests in the current study suggests that the items were not retrieved via episodic LTM processes during the original WM task. Here, we used a task that is essentially identical to the double-retrocue tasks that have shown the return to baseline pattern of decoding of the uncued item in several studies (Lewis-Peacock et al., 2012; LaRocque et al., 2014; Rose et al., 2016). If the uncued item was reactivated with episodic LTM retrieval processes when subsequently cued as the target item (for Probe 2 switch trials/UA items), subsequent LTM should be greater for these items compared with items maintained in and reported directly from focal attention (AU items). However, there was no difference between these items in LTM. Thus, it is more likely that LTM was not differentially required depending on the state in which items were retained in WM.

The findings from two additional analyses provide support for this interpretation. First, there were no significant disparities in LTM performance between items that were either correct or incorrect on the WM task (see Supplementary Figure 1). This means that subsequent LTM performance was similar regardless of whether the items were successfully maintained and retrieved from WM. If an item was incorrect on the initial WM task, it implies that it was not held in the focus of attention throughout the trial. In contrast, a correct item could have been maintained and retrieved successfully in the WM task either because it remained in the focus of attention throughout the trial or because it was initially dropped from focal attention but subsequently retrieved, possibly via episodic LTM retrieval processes. It would be reasonable to assume that items correctly maintained and retrieved in the WM task would have a higher likelihood of being successfully retrieved during the LTM tests, especially if they were initially retrieved through episodic LTM retrieval processes due to the benefits of retrieval practice on LTM. However, this was not observed. Hence, the absence of a distinction in subsequent LTM performance between correct and incorrect items from the WM task suggests that there was no advantage to LTM for items that were differentially processed within the focus of attention during the initial WM task.

Second, there was no correlation between the difference between Probe 1 versus Probe 2 switch trials on WM and subsequent LTM for the critical UA versus AU items (see Supplementary Table 3). That is, participants who showed a greater effect on WM, which suggests that they dropped the uncued item from focal attention and then successfully reactivated it, showed similar subsequent LTM to those who did not exhibit as much of an effect. In other words, the size of the switch difference in WM did not predict the likelihood of retrieving the unattended items from LTM later on, as would be predicted if returning uncued items to focal attention requires episodic LTM retrieval.

Based on a single-item focus model, it was proposed that these "latent" items were stored and reactivated through LTM processes (McElree, 2006). Therefore, it was expected that subsequent LTM performance would be superior for items that were reactivated during switch trials (UA items) compared with items that were initially held in focused attention, dropped, and not reactivated during the WM task trial (AU items). However, our findings did not provide any 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 ("activity-silent"?) state, ready to be used for ongoing processing. The proposed region of direct access has a capacity limit of 3–4 items or "chunks." The current double-retrocue WM task only required retaining two items per trial, so the capacity was not exceeded. The latent item may stay 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. If the capacity of the region of direct access was exceeded, then switching attention to recover passively retained items outside of focal attention would implicate LTM processes and affect subsequent LTM, according to this model. Young adults showed no switch cost relative to Probe 1 and an advantage of staying on the same probe, whereas older adults showed a switch cost relative to both Probe 1 and Probe 2 stay. The latter might suggest an LTM deficit in older adults compared with young adults, but there was no influence of item type or age on LTM performance. Consequently, the results imply that switching between two items in WM within the current paradigm does not implicate LTM. The inability to decode an unattended item observed in prior research must be attributed to factors other than retrieval from LTM.[8]

The embedded processes model might also be able to explain the results of the current study. Cowan (1988, 2019) suggested that perception, attention, and LTM processes are all embedded processes in WM. Items that are perceived and are continuously attended are activated in the focus of attention; when attention shifts away from the item, it may remain cognitively "activated" in the "activated portion of long-term memory" (aLTM). Rose et al. (2016) suggested that the term "prioritised LTM" may be more appropriate than aLTM, as there is a lack of sustained, elevated neural activation for such passively retained latent items. Nonetheless, the model may be seen to accommodate such passively retained prioritised items if it incorporates a distinction between the retrieval processes associated with reactivating passively retained, latent items and episodic LTM retrieval processes that result in superior subsequent LTM (Cowan, 2008, 2019).

If latent items in WM cannot be retained through continuous active processing or retrieval from LTM, then what cognitive processes are responsible for their retention? Attentional refreshing is a hypothesised maintenance process that is proposed to be distinct from both rehearsal and episodic retrieval (Camos & Barrouillet, 2011; M. K. Johnson, 1992). Recently processed items that are outside of the current focus of attention, but are still relevant for ongoing cognition, may be periodically refreshed by bringing them back into the focus of attention to retain them until they are no longer needed (Camos et al., 2018).

What neurobiological evidence supports the short-term retention of information that is neither actively held in focal attention nor passively stored in LTM? The synaptic theory of WM suggests that short-term synaptic-plasticity mechanisms might be responsible for retaining such items (Mongillo et al., 2008; Stokes, 2015; Trübutschek et al., 2017). Actively retaining an item in WM is associated with sustained, elevated neural activity for the item. When 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 refresh the item. In this case, reactivating deprioritised items might not need LTM retrieval.

Moreover, this putative refreshing process is hypothesised to decline as people age (Gaillard et al., 2011). The age differences in WM on Probe 2 switch trials, alongside the lack of an age difference on Probe 1 trials and on the LTM tests, may be seen to provide support for this hypothesis. Furthermore, refreshing can be disrupted by the inclusion of an attention-demanding secondary task (e.g., Barrouillet et al., 2007). That dividing young adults' attention disrupted their performance may also be seen to provide support for this hypothesis. Most importantly for present purposes, neither age nor dividing attention during maintenance had any systematic effects on subsequent LTM (except for high-confidence location memory for older adults and item memory for young adults with DA). This suggests that whatever maintenance process was involved to help retain the items on the WM task, it appears to be distinct from episodic LTM retrieval.

The lack of an effect of the retrocue condition (AA, AU, UA, UU) on subsequent LTM is interesting compared with the previous literature that used retrocues in WM tasks and tested the effect on subsequent LTM in young and older adults. As reviewed in the introduction, Strunk et al. (2019) found that older and young adults both benefitted from single retrocues for WM and LTM, but only for item memory, not location memory. It may seem surprising that older adults in both Strunk et al. (2019) and the current study did not have deficient location or associative LTM overall for items initially maintained in WM. It is important to note that the older adults in this sample did show LTM deficits on both the Immediate and Final Free Recall LTM tests of the TICS neuropsychological screening test. So, this sample of older adults was representative of the population in that regard. We suspect that the lack of an age-related deficit on the location and associative recognition test was due to the older adults being able to rely on the familiarity signal evoked by the "old" stimulus relative to the "new" stimulus on these two-alternative forced-choice tests. This may be because the item-location bindings were only temporarily maintained in WM and were actively deleted from WM at the end of each trial—that is, they were not strongly consolidated in LTM (Oberauer, 2005; Oberauer & Werner, 2022). Future studies could test this hypothesis with subsequent associative LTM recognition tests that present all old items (for both targets and lures) and ask participants to recall which one was the associated item (intact vs. scrambled pairs).

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 reactivation effect (stay vs switch) for both the young and old groups on the WM task. However, the retrocue benefit is slightly different from our previous studies (LaRocque et al., 2015; Rose et al., 2016), which showed a significant difference between Probe 1 and Probe 2 switch responses. The main difference between the paradigms is that, because we were worried that subsequent LTM could be at chance levels of performance for the older adults and young adults with DA, we warned participants that their LTM for the items, their location, and their associated pair would be tested at the end of the session. In previous studies (on healthy young adults with FA), the participants were not forewarned about an upcoming LTM test of the items from the WM task (LaRocque et al., 2015; Rose et al., 2016). This may have caused our participants to engage in deeper, more elaborative encoding of the items and their associated location and pair than in previous studies, which may have affected WM performance. It was also possible that the intentional LTM encoding of items during the WM task may have engendered additional elaborative strategies that reduced the requirement of retrieval from LTM for all items, including the critical control (AU) items.

To test for this possibility, the third control experiment was conducted. It is also noteworthy that the pattern of subsequent LTM for UA versus AU items is similar between Experiment 3 with "incidental encoding" and Experiment 1 with "intentional encoding" for the subsequent LTM test. Experiment 3 also replicated the results of LaRocque et al. (2015), with no difference between AU and UA items of subsequent LTM test. Moreover, Rose et al. (2016, Experiment 1) tested participants' subsequent LTM for the items maintained during the initial WM task performed in the fMRI scanner on a subsequent old/new item recognition test and also found no difference between UA and AU items (see Supplementary Figure 5, Rose et al., 2016). Indeed, in all of these studies, the results suggested that reactivating latent items does not involve retrieving them using episodic LTM retrieval processes. That is, the level of subsequent recognition performance is similar between studies with incidental and intentional encoding. Therefore, these data help rule out the possibility that participants engaged in some kind of elaborative processing, for example, during the inter-trial interval, because they were anticipating the upcoming LTM test.

With regard to how the present results relate to those of LaRocque et al. (2015), we replicate and extend the finding that subsequent LTM is largely unaffected by retaining the items in different WM states during the two-item double-retrocue task. This finding was replicated despite differences in the participants (older adults, young adults with DA), and aspects of both the WM task (stimuli, timing, intentional encoding) and the LTM tests (more detailed assessment of item, location, and associative LTM, as well as participants' confidence in these judgements).

On the nature of age differences on WM/LTM. The current study is part of an overarching project on the nature of age differences in WM, particularly with regard to older adults' ability to use attentional control processes to proactively prioritise the active maintenance of the cued item in focal attention as well as passively retain an uncued item outside focal attention and then reactivate it when it is subsequently cued. In relation to age differences on WM, a critical question emerging from prior research is whether or not older adults can use retrocues to prioritise cued items, and retain and reactivate uncued items back into focal attention as well as young adults.

Most prior research examining age effects on retrocue benefits used tasks with a single retrocue. Loaiza and Souza (2018) used 0, 1, or 2 retrocues and a recall task to investigate age difference in the ability to reactivate latent WM items. They found that older adults obtained a similar benefit of retrocues as young adults, but, in a follow-up study, older adults did not retain a retrocue benefit as cognitive load increased (Loaiza & Souza, 2019). Our double-retrocue recognition task revealed that the way the older adults used the retrocues appeared to differ from young adults. Relative to Probe 1, young adults' accuracy improved on cue-stay trials, which is sensible given that the same item was cued and tested on both Probe 1 and 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, relative to Probe 1 accuracy, they showed a decrease in accuracy on Probe 2 switch trials. This suggested that older adults were not using 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/non-match decision based on the strength of the familiarity signal. We (Xu et al., in press) recently proposed and tested a dual mechanism of control account of age differences in WM that explains the variability in age differences in reactivating latent items seen in this study and the prior literature. This account proposes that age differences in WM depend on the extent to which the task situation either requires older adults to engage in proactive maintenance processes or allows them to rely on a reactive, familiarity-based mode of processing.

With regard to the nature of age differences and the role of LTM in WM, two studies examined young and older adults on both WM and subsequent LTM tests of the items. Bartsch et al. (2019) matched young and older adults' performance on a WM binding task for word pairs (by extending the amount of encoding time for older adults), yet subsequent associative LTM for the items revealed an age deficit. Forsberg et al. (2022) matched young and older adults encoding time on a WM task with visual icons of objects and found deficits on both WM and subsequent LTM for the objects, but there was no age difference in the ratio of WM items that were forgotten on a subsequent (item recognition) LTM test, similar to our results. An important difference with the Bartsch et al. (2019) study is that they used verbal stimuli, which may have implicated different maintenance, elaboration, or retrieval processes than our study, which used both visual and verbal stimuli in a double-retrocue task. We found that there were age deficits in WM (and the difference was not larger for reactivating the latent items), and subsequent LTM, but only for high-confidence location memory. Collectively, these findings highlight that, to fully capture the size and type of age differences in WM and subsequent LTM, it is important to assess young and older adults on different types of WM (e.g., with and without matched encoding times, verbal and non-verbalisable visual stimuli) and subsequent LTM tests (e.g., those that assess associative binding in addition to item memory). We argue that such an approach is essential for assessing the full picture with regard to the roles of attentional control processes and LTM retrieval processes in age differences in WM. Based on the full pattern of results in this and prior studies (Bartsch et al., 2019; Forsberg et al., 2022), as well as our related study and dual mechanisms of control account of age difference in WM (Xu et al., in press), we suggest that older adults' deficits on this and other retrocue WM tasks are largely attributable to deficiencies in attentional control processes—not episodic LTM processes.

Caveats, potential limitations, and future directions. The lack of effects of the factors of age and DA on the LTM tests may raise concerns about the sensitivity of the measures. Effects of age and DA on episodic LTM (especially associative memory) are common, and the absence of either effect may suggest that the LTM test was not sensitive to the experimental manipulations. However, the lack of an age effect on associative memory is likely because the lure stimulus on each trial of the two-alternative forced-choice test was novel. Therefore, participants could likely select the correct item based on whichever item felt more familiar. Furthermore, the LTM test performance *was* sensitive to other commonly observed effects, such as the testing effect, such that the accuracy of AA items was greater than that of UU items. Although having novel lures on the associative LTM test limits its ability to assess recollection of the associated item, several lines of evidence attest to the robustness of the effects that were observed. We assessed the effects of reactivating latent WM, age, and DA, on item-, location-, and associative-recognition tests for responses made with either high or low confidence for items that were either correctly or incorrectly recalled on the initial WM task. None of these manipulations showed better memory for the critical comparison between the UA and AU items. That is, every opportunity to observe the effect failed to support the LTM account of latent WM even without correction for the multiple comparisons. Moreover, recall that LaRocque et al. (2015, Experiments 2 and 3) and Rose et al. (2016, Experiment 1) also found no evidence that reactivating a latent WM item benefitted its subsequent LTM. Therefore, despite the multiple means by which the effects could have been shown, all the evidence counters the notion that retrieval from LTM is required to retain and reactivate a latent item on the double-retrocue WM task. Nonetheless, future research on this topic should include tests that are both sensitive to age differences in associative LTM (e.g., recall tests) and simultaneously avoid floor and ceiling effects on both WM and subsequent associative memory across groups of young and older adults.

With regard to the potential concern that we only tested WM for set sizes of two items per trial, substantial research suggests that adults have the capacity to actively maintain up to three or four items in the focus of attention in WM (Cowan, 2001, 2011), yet others argue that only a single item or chunk may be retained in focal attention (McElree, 2006; Oberauer, 2009). Although our experiments only tested two items per trial, it is unlikely that participants actively maintained both items in focal attention throughout the trial. Several studies with the same paradigm (Lewis-Peacock et al., 2012; LaRocque et al., 2014; Rose et al., 2016; Wolff et al., 2017) have shown the return-to-baseline levels of decoding for an uncued item with only two memory items per trial. This suggests that participants do not actively maintain an uncued, latent item in a sustained manner in WM; yet it can be reactivated when attention is shifted back to them (LaRocque et al., 2014; Lewis-Peacock et al., 2012; Rose et al., 2016), or by a pulse of TMS (Rose et al., 2016) or by a nonspecific visual (Wolff et al., 2017) or auditory (Mamashli et al., 2021; Wolff et al., 2020) impulse stimulus. This was certainly the case in Experiment 2 when participants had to perform the DA task during the delays; the uncued item was likely dropped from focal attention due to the demanding nature of the DA task, yet successfully reactivating this latent WM did not result in superior subsequent LTM compared with control items. Therefore, although it is difficult to know for certain that the uncued items in this study were dropped from focal attention, passively retained and reactivated as in our previous studies using the same paradigm (and related studies), especially since we did not record neuroimaging data and apply delay-period TMS or sensory impulses, there is no compelling reason to believe that the situation in this study would be any different. Nevertheless, future studies should include neuroimaging, brain decoding, and brain stimulation techniques to investigate the role of LTM in WM and the effects of reactivating latent items, ageing, and DA on WM and subsequent LTM. Studying these topics can help the field to better understand the relationship between WM and LTM more generally. 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 preregistered report predicted that if retaining and reactivating latent items in WM involves episodic LTM processes, then requiring participants to switch attention away from an item and reactivate it should affect subsequent LTM. However, the results showed that, although switching away from and reactivating uncued items affected WM, as did the effects of age and DA, none of these factors had consistent, reliable effects on LTM. Thus, the results are inconsistent with the hypothesis that latent WM items are retrieved via LTM processes at least for this double-retrocue task in which there were only two items to remember per trial and interference was minimal. Instead, the various results collectively provide further evidence that latent WM items may be retained via short-term synaptic plasticity mechanisms. These neurobiological retention mechanisms may be reflected by the cognitive concepts described as the attentional refreshing of items held outside the focus of attention in either a region of direct access or the activated portion of LTM, although it is important not to conflate cognitive activation with neural activation.

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. This ability is one of the most intriguing areas of research on WM and LTM, yet it remains relatively poorly understood. Modern cognitive, neural, and computational 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.

Acknowledgements

The authors acknowledge Bernadette Widjawa, Lauren Crowe, and Geralynn Timm for their assistance with data collection, and Joseph Saito, Nicholas Yeh, and Morgan Widhalm for their help with an initial pilot study. They also thank Nelson Cowan for suggesting the third experiment and for He and Eren Günseli’s helpful comments on an earlier version of this manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the National Science Foundation’s CAREER Grant #1848440 awarded to N.S.R.

ORCID iDs

Chang Mao Chao https://orcid.org/0000-0002-4087-6176
Vanessa Loaiza https://orcid.org/0000-0002-5000-7089

Data accessibility statement

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/ztqx8/

Supplementary material

The supplementary material is available at: qjep.sagepub.com

Notes

1. It is important to consider encoding processes. There is recent discussion about a potential difference between "WM encoding" and "LTM encoding" (Rose & Chao, 2022; Slotnick, 2022). In our view, as soon as a stimulus is perceived, there will be a trace that is represented and consolidated in memory irrespective of whether one is performing a WM or LTM test (or no memory test at all). Such representations can be actively maintained in WM and subsequently accessed and influence performance on LTM tests. As discussed further on, we argue that it is inaccurate to distinguish between "WM encoding" and "LTM encoding", e.g., suggesting that participants could encode the uncued item "into LTM" instead of WM. From our perspective, such a distinction does not make sense. While it is true that longer versus shorter presentation times allow for more deep/conceptual/elaborative processing, which could result in the memory representations being more accessible on LTM tests, this should not be taken to mean that such items are encoded ''in LTM" and not "in WM".
2. Note that it is not yet clear what mechanisms underlie "activity-silent" maintenance. The term "activity-silent" refers to situations where the same decoding method that could detect active maintenance of items in the focus of attention in WM can no longer decode an item that has been dropped from focal attention. There are a number of plausible candidates that could explain how "activity-silent" WM maintenance might occur (for review, see Stokes, Muhle-Karbe, & Myers, 2020). It could be that there is active representation of deprioritized information retained in WM, but the representation is reduced (subthreshold) or even suppressed below baseline such that it is not detectable by the decoding method (Schneegans & Bays, 2017); or it could be that the information is actively maintained, but is represented by either elevated neural activity that is either in an 'orthogonal' state or in different, unmeasured neural populations (Christophel, loshchinina, Allefeld & Haynes, 2018). In contrast to the LTM account, these candidate mechanisms that could give rise to activity-silent WM are all short-term retention mechanisms - not LTM processes per se.
3. They refused to remain seated in the instructed position and distance from the screen and made random responses as fast as possible to finish the experiment as quickly as possible.
4. 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.
5. Note that, although the lack of an age difference on the location or associative LTM tests may be surprising, the older adults in this sample did show LTM deficits on both the Immediate ($p<.001$) and Final Free ($p<.0001$) Recall Tests of the Telephone Interview of Cognitive Status. As discussed further on, we suspect that the lack of age-related deficits on the associative recognition test was due to the older adults being able to rely on the familiarity signal evoked by the 'old' stimulus relative to the 'new' stimulus on the two-alternative forced-choice test.
6. Note that, because the UA items were not tested on probe-1 and AU items were not tested on probe-2, the effect size for their difference on WM could not be calculated. This effect size comparison was recomputed to compare AA to UA items on both WM and LTM. The same general pattern was observed (see Supplemental Figure 6).
7. The exceptions are that older adults were worse than the young FA group on high-confidence location memory responses and young adults with DA had better high-confidence item memory for U-A than A-U items. Importantly, we were able to fairly compare subsequent LTM for the items (i.e., forgetting rates) by focusing analyses on items that were initially remembered correctly on the WM task.
8. Note that it is unclear if activity-silent latent WM representations contribute to the capacity limit of the region of direct access. If so, then there must be a limit to the number of activity silent representations that can be retained in the region of direct access, and if this number is exceeded then reactivating latent WM items must involve LTM processes. Evidence of proactive interference and serial dependence biases from no-longer-relevant "deleted" WM representations on recall of the currently relevant target item may be seen to argue against the notion of activity-silent latent WM representations being represented in the region of direct access. However, studies that have shown that no-longer-relevant deleted representations cannot be reactivated with TMS or influence behavior suggest that passively retained latent items can be actively deleted from WM (Rose et al., 2016; Fulvio & Postle, 2020), possibly from the region of direct access. Rather, biases from no-longer-relevant items, especially those from set sizes that exceed the capacity of the region of direct access, are more likely an influence from LTM representations. Whether or not there is a capacity limit for activity-silent latent WM, and the distinction between items retained in the region of direct access vs. LTM needs to be clarified in future research.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, *2*, 89–195.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

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), 570–585.

Bartsch, L. M., Loaiza, V. M., & Oberauer, K. (2019). Does limited working memory capacity underlie age differences in associative long-term memory? *Psychology and Aging*, *34*(2), 268–281. https://doi.org/10.1037/pag0000317

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.

Beukers, A. O., Buschman, T. J., Cohen, J. D., & Norman, K. A. (2021). Is activity silent working memory simply episodic memory? *Trends in Cognitive Science*, *25*(4), 284–293. https://doi.org/10.1016/j.tics.2021.01.003

Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: From passive maintenance to active refreshing. *Developmental Psychology*, *47*(3), 898–904.

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), 19–32.

Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nature Neuroscience*, *21*(4), 494–496.

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

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford University Press.

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 114–185. https://doi.org/10.1017/S0140525X01003922

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

Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia*, *49*(6), 1401–1406.

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

Craik, F. I. M. (1986). A functional account of age differences in memory. *Human Memory and Cognitive Capabilities: Mechanisms and Performances*, *5*, 409–422.

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

Craik, F. I. M., & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1729–1739.

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.

Duarte, A., & Dulas, M. R. (2020). Episodic memory decline in aging. In A. K. Thomas & A. Gutchess (Eds.), *The Cambridge handbook of cognitive aging: A life course perspective* (pp. 200–217). Cambridge University Press.

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.

Forsberg, A., Guitard, D., Greene, N. R., Naveh-Benjamin, M., & Cowan, N. (2022). The proportion of working memory items recoverable from long-term memory remains fixed despite adult aging. *Psychology and Aging*, *37*, 777–786.

Foster, J., Vogel, E., & Awh, E. (2019). Working memory as persistent neural activity. In M. J. Kahana & A. D. Wagner (Eds.), *Oxford handbook of human memory*. Oxford University Press.

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. 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.

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.

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

Jeneson, A., & Squire, L. R. (2012). Working memory, long-term memory, and medial temporal lobe function. *Learning & Memory*, *19*(1), 15–25.

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, *4*(3), 268–280. https://doi.org/10.1162/jocn.1992.4.3.268

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*(4), 371–376.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annu. Rev. Psychol*, *59*, 193-224.

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., 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.

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*, Article 5.

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., & Camos, V. (2016). Does controlling for temporal parameters change the levels-of-processing effect in working memory? *Advances in Cognitive Psychology*, *12*(1), 2–9.

Loaiza, V. M., & McCabe, D. P. (2012). Temporal -contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, *40*, 191–203.

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., & 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., & Souza, A. S. (2019). An age-related deficit in preserving the benefits of attention in working memory. *Psychology and Aging*, *34*(2), 282–293.

Mamashli, F., Khan, S., Hämäläinen, M., Jas, M., Raij, T., Stufflebeam, S. M., Nummenmaa, A., & Ahveninen, J. (2021). Synchronization patterns reveal neuronal coding of working memory content. *Cell Reports*, *36*(8), 109566.

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.

McElree, B. (2006). Accessing recent events. *Psychology of Learning and Motivation—Advances in Research and Theory*, *46*(6), 155–200. https://doi.org/10.1016/S0079-7421(06)46005-9

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.

Morris, R. G., Craik, F. I., & Gick, M. L. (1990). Age differences in working memory tasks: The role of secondary memory and the central executive system. *The Quarterly Journal of Experimental Psychology*, *42*(1), 67–86.

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

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.

Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U., & Nilsson, L. G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *51*(4), P234–P240.

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), 368–387.

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45–100.

Oberauer, K., & Awh, E. (2022). Is there an activity-silent working memory? *Journal of Cognitive Neuroscience*, *34*, 2360–2374.

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. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958.

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

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*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y

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

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., & Chao, C. M. (2022). Hippocampal involvement in working memory following refreshing. *Cognitive Neuroscience*, *13*, 215–217.

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), 1019–1029.

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., Myerson, J., Roediger, H. L., III, & 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), 471–483.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

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

Schneegans, S., & Bays, P. M. (2017). Restoration of fMRI decodability does not imply latent working memory states. *Journal of Cognitive Neuroscience*, *29*(12), 1977–1994.

Singh, B., Gambrell, A., & Correll, J. (2022). Face templates for the Chicago face database. *Behavior Research Methods*, *55*, 639–645.

Slotnick, S. D. (2022). The hippocampus and long-term memory. *Cognitive Neuroscience*, *13*, 113–114.

Souza, A. S. (2016). No age deficits in the ability to use attention to improve visual working memory. *Psychology and Aging*, *31*(5), 456–470.

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., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory: Retro-cues enhance retrieval and protect from visual interference. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 890–910.

Stokes, M. G. (2015). “Activity-silent” working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405.

Stokes, M. G., Muhle-Karbe, P. S., & Myers, N. E. (2020). Theoretical distinction between functional states in working memory and their corresponding neural states. *Visual Cognition*, *28*(5–8), 420–432.

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*, Article e23871.

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), 104–132.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*, 813–826.

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.

Wolff, M. J., Kandemir, G., Stokes, M. G., & Akyürek, E. G. (2020). Unimodal and bimodal access to sensory working memories by auditory and visual impulses. *Journal of Neuroscience*, *40*(3), 671–681.

Xu, C., Chao, C. M., & Rose, N. S. (in press). A Dual Mechanisms of Control Account of Age Differences in Working Memory. Psychology & Aging

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.