



# Does unitization really function like items? The role of interference on item and associative memory processes

Catherine M. Carpenter<sup>1</sup> · Nancy A. Dennis<sup>1</sup>

Accepted: 20 December 2022  
© The Psychonomic Society, Inc. 2023

## Abstract

Associative memory declines as we age, while item memory remains relatively stable. Previous work has shown that, in both younger and older adults, while item recognition declines linearly across time and interference, associative recognition declines only with longer delays and more interference. Unitization is a memory process found to support associative memory by allowing pairs to presumably be processed like single items. Research has found that unitization can benefit memory in aging by boosting associative memory to be on par with that of younger adults. Yet it remains unclear exactly the mechanism responsible for this enhancement in memory. The current studies aimed to determine whether unitized pairs show similar memory to that of items or associations with increasing time and interference, and determine how physically similar unitized pairs must be to perform like items and examine the effect of age on unitization in a continuous recognition paradigm. The results show that while unitized pairs exhibit higher corrected recognition compared with associative pairs at all lags, unitized pairs are not remembered to the degree that items are. It is critical that unitization boosts accurate recognition of pairs in both age groups across all early and middle lags compared with associative pairs. The results suggest that unitization may promote a more efficient associative link than unrelated associations over increasing time and interference, but the benefit does not reach that of item memory. These results demonstrate that while unitization benefits corrected recognition with earlier interference, its effect may not hold with later interference.

**Keywords** unitization · associative memory · aging · continuous recognition

## Introduction

The associative deficit in aging describes the often-observed pattern of memory deficits whereby memory for the association amongst pairs of items tends to show significantly larger age-related declines compared with memory for single items (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). For example, previous research has suggested that while older adults may be able to remember a single word (or a list of words) as well as a younger adult, they have more difficulty binding together words as an associative pair in memory as compared with younger adults (Naveh-Benjamin, 2000). The associative memory deficit in aging has been observed for face–name (Naveh-Benjamin

et al., 2004), emotional–word–gender (Naveh-Benjamin et al., 2012), item–context (Hennessey et al., 2018), and object–object associations (Saverino et al., 2016). This deficit in remembering associative information is believed to lie in a decreased ability to bind information together during episodic encoding (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). Because of the severity of issues that may result from faulty associative memory (ranging from embarrassment in failing to remember face–name associations, to health issues when failing to remember medication–dosage associations) it is important to understand exactly why these associative deficits occur as well as find ways to mitigate age differences.

In addition to these age deficits in associative memory, retention intervals and interference can also affect both item and associative memory. Historically, research has found that increased time between study and test, results in a decrease in memory, one paramount experiment being that of Ebbinghaus’s forgetting curve (Roediger, 1985) and more recent work using a continuous recognition paradigm

---

✉ Nancy A. Dennis  
nad12@psu.edu

<sup>1</sup> The Pennsylvania State University, Psychology,  
State College, PA, USA

(Ashford et al., 2011; Berman et al., 1991; Jones & Atchley, 2002; Kuhlmann et al., 2021; Poon & Fozard, 1980). The continuous recognition paradigm features a constant encoding and simultaneous retrieval of information, such that there is no break between the two tasks, with new and repeating information being intermixed in a continuous stream (Hockley, 1982). Using the continuous recognition paradigm, research has found that longer retention intervals, or lags, tend to produce worse memory performance compared with shorter retention intervals due to interference from related information (Chen & Naveh-Benjamin, 2012; Kim et al., 2001; Portrat et al., 2008). In this literature, “lag” refers to the number of trials between the first presentation of a trial and the corresponding target or lure trial later in the information stream (Kuhlmann et al., 2021). Interference and delay are operationalized in this work (and the current task) by utilizing the lag manipulation in order to create distance, and introducing interfering trials, between study and test trials. Thus, longer lags incur greater distance and more interference in memory compared with shorter lags (Chen & Naveh-Benjamin, 2012; Kuhlmann et al., 2021).

While earlier work suggests that delay and interference have similar effects on associative memory and item memory across the life span with respects to shorter term memory for pictorial stimuli (e.g., lags up to 10 seconds; Chen & Naveh-Benjamin, 2012), results from recent work suggest that item and associative memory for semantic stimuli may be differentially affected by interval length (Kuhlmann et al., 2021). Specifically, Kuhlmann et al. (2021) found that while both item and associative memory declined across lags of varying lengths, item memory declined linearly across all lags, while associative memory remained relatively stable at shorter lags (1 and 11 trials), declining only in later lags (24 and 44 trials). While age did not interact with lag, there was a significant interaction between age and type of memory, such that older adults performed worse on the associative memory trials than younger adults, demonstrating the typical associative deficit in aging. The authors interpreted the difference between item and associative declines across lags as associative memory having greater resistance to interference than item memory (Kuhlmann et al., 2021). These results are consistent with both the associative deficit hypothesis (Naveh-Benjamin, 2000) and the memory-system dependent forgetting hypothesis, which indicates that interference should not have as large of an effect on associative information as it will on item memory due to efficient pattern separation in the hippocampus, a region in which associative memory is processed (Hardt et al., 2013).

Despite the differences observed between item and associative memory, there is a vein of research that focuses on how associative memory may mirror item memory, through means of unitization. Unitization is the process by which one can create a meaningful connection between individual

items in order to create a single, bound representation (Diana et al., 2008; Ford et al., 2010; Giovanello et al., 2006; Graf & Schacter, 1989). One example of this would be the words *mail* and *box*. Separately, they are two individual words with meanings of their own, but if one wanted to remember them together as an associative pair, forming the compound word *mailbox* is meaningful because most individuals recognize it as its own word. Research has also shown that unitization at the time of encoding improves both implicit and explicit memory compared with nonunitized associative encoding (Bader et al., 2010; Diana et al., 2008; Graf & Schacter, 1989; Quamme et al., 2007).

Unitization has also been found to be supported by the use of familiarity at retrieval (Diana et al., 2008), and is reliant on cortical MTL regions (e.g., perirhinal cortex), making it more similar to item memory, which is reliant on similar regions, than hippocampal-based recollection of associative information (Delhay & Bastin, 2018). To this end, previous work in the domain of aging has shown that unitization applied to associative memory tasks can help older adults overcome age-related associative memory deficits (Bastin et al., 2013), as they utilized item-based familiarity processing during memory retrieval. Work with compound word pairs acting as the unitized condition have found similar results such that older adults performed similarly to younger adults in the unitized condition, while older adults still exhibit an associative deficit for the nonunitized associative condition (Ahmad et al., 2014; Delhay & Bastin, 2018). However, it is unclear whether unitized pairs truly operate at the level of items and whether this would be maintained in healthy aging. The current set of studies aimed to examine the effects of interference and retention intervals on unitized pairs, nonunitized pairs, and items as measured by a continuous recognition task that manipulated lag length, replicating and expanding upon the work of Kuhlmann et al. (2021). As such, we felt that a continuous recognition task with lags introduced would be more suitable to examine the effects of interference and retention interval as measured by lag on recognition performance of unitized, nonunitized pairs, and items.

## Experiment 1

The aim of Experiment 1 was to investigate whether unitization during encoding would result in associative memory for words pairs that equaled that of item with respect to memory or if it is simply an efficient way to boost item–item associative memory. Building on the work by Kuhlmann et al. (2021), this experiment was designed to test whether lag delays and interference act similarly on unitized pairs as it does for items, or whether unitized pairs will act as associative pairs do. It was hypothesized that if unitization operates to create a single

representation of individual items, then memory for unitized pairs should decrease linearly across all lags, in a manner mirroring item memory. Additionally, it was expected there to be no age differences in memory performance across lags in the unitization condition. Alternatively, if unitization is simply a more effective way to promote associative binding, then it would be expected for memory for unitized pairs to mirror that of associative pairs, remaining stable until later lags where performance declines would emerge. This experiment was also interested in examining whether older adults can unitize word pairs in the same manner as younger adults or if age differences exist in the mechanism underlying unitization.

## Methods

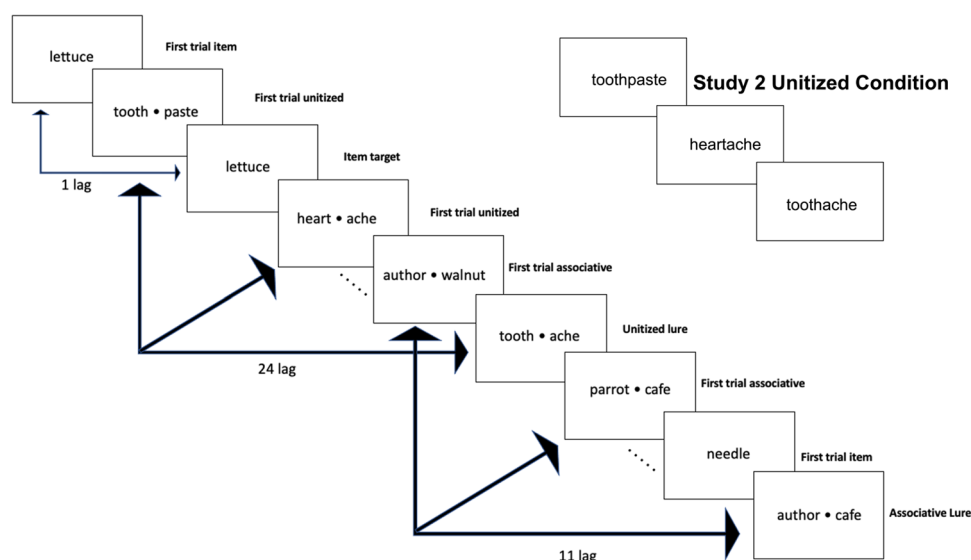
### Participants

A power analysis, based upon the sample size of Kuhlmann et al. (2021), showed that in order to reach a medium effect size of  $f = .25$ , 55 participants in each age group would be needed. Following, 58 younger adult participants were recruited through the psychology subject pool at Penn State University, where they participated in-person at the Pennsylvania State University and received course credit for their effort. Fifty-eight older adult participants completed the tasks online through Prolific and were paid \$4.72 for participating. Participants were overrecruited in order to reach the suggested  $n$ , however some subjects warranted removal following their participation. Participants with higher than a 10% no-response rate (two young), performance at chance on any of the given conditions (one young; four old), or participants with lost data (one young) were removed from the data set. Thus, 54 younger adults ( $M_{\text{age}} = 18.60$ ,  $SD_{\text{age}} = 2.74$ ; 46 female;  $M_{\text{education}} = 13.32$  years,  $SD_{\text{education}} = 0.67$ ) and 54 older adults ( $M_{\text{age}} = 65.35$ ,  $SD_{\text{age}} = 4.54$ ; 24 female;  $M_{\text{education}} = 16.41$  years,  $SD_{\text{education}} = 2.29$ ) were included in the analyses.<sup>1</sup> Younger adult participants identified as White (39), Asian (4), Black or African American (4), Asian/Native Hawaiian/Other Pacific Islander (2), or more than one race/other (5). Older adult participants identified as white (51), Black or African American (2), or Asian (1). Participants provided a waiver of documentation of consent for a protocol approved by the Pennsylvania State University Institutional Review Board. Demographics were collected prior to beginning the task. Both online and in-person testing followed the same protocol, with instructions printed on the screen. Following the task, participants were debriefed.

<sup>1</sup> Older adult participants were tested online and were rejected for high numbers of no-response rates and failure to complete the task; these counts are not reflected in the total number of “removed” participants.

### Stimuli and design

The stimuli consisted of 48 single words, 48 associative word pairs (word–word associations), 72 unitized word pairs (unitized associations), and 48 foils (novel words and word pairs). A compound word list used in previous studies of unitization was used for selecting the unitized word pairs. This list contains compound words in sets of three (two unrelated compound words and one compound word made from those two words) as to allow for recombined lures (Ford et al., 2010). For example, *light • weight* and *club • house* were initially presented, and *light • house* would be the recombined lure. Word sets were derived from a previously used set of compound words examining unitization (Ford et al., 2010). Associative word pairs were taken from Kuhlmann and colleagues recent work (Kuhlmann et al., 2021), and random unrelated words were used as items (see Supplemental Material for complete word list). The unitized and associative word pairs were presented separately with a dot and two spaces on either side between the two components of the pair to delineate a space between the two words. These words were normed during piloting of the task in a younger adult population, in which a small group of participants were asked if there was any repetition or similarity in the words or pairs. The task was created and run through PsychoPy and Pavlovia, respectively. The words and word pairs were presented to participants across two runs, for a total of 180 trials in each run (24 items, 30 associative pairs, and 72 unitized pairs initially seen in each run prior to the second presentation whether that be a lure or target trial. The associative and unitized word pairs had higher numbers due to the constraints of the continuous recognition paradigm and the rearrangement of lure pairings). Stimuli were presented one at a time (either a pair or one item) and remained on the screen for 5 seconds. Half of the trials were later re-presented as targets (i.e., the same item/pair presented a second time) at one of the four lags (i.e., 1, 11, 24, 44) to mirror that of Kuhlmann et al. (2021). The other half of the trials were followed by a lure in the same follow-up lag position (there were three targets and lures in each memory type condition per lag in each run). Lags refer to the number of trials in between the current trial and the corresponding lure or target trial later in the task. For single words, a lure was a brand-new word (unrelated lure/foil) in place of the original word, and for word pairs a lure was a recombined associative pair from a pair presented zero to four intervening trials away (see Fig. 1). Since the lags needed to be exact to be consistent with the demands of the continuous recognition paradigm, the trials were pseudorandomized in each run such that there were no more than three consecutive target and lure test trials of any one memory type. Each trial was shown in Arial font and was centered in the screen. The letter height of each trial in PsychoPy was 0.1.



**Fig. 1** Task paradigm. The only difference from Study 1 and Study 2 is the unitized condition in which the spacing was removed between the two words in the pair to create a compound word

## Procedure

The procedure was similar to that of Kuhlmann et al. (2021). Participants read written instructions on the screen prior to beginning the task. The instructions explained that participants would view words and word pairs and would be asked to indicate whether they have seen that word or pair previously. They were instructed to indicate “yes” if the word or pair was old meaning they remember seeing the word or pair together previously, or to indicate “no” if the word or pair was new to them meaning they had not seen the word or pair together previously. Participants then completed 10 practice trials prior to beginning the actual task. The task itself, excluding the instructions, took participants approximately 25 minutes to complete. Participants were debriefed following the task (see Fig. 1 for paradigm design).<sup>2</sup>

## Results

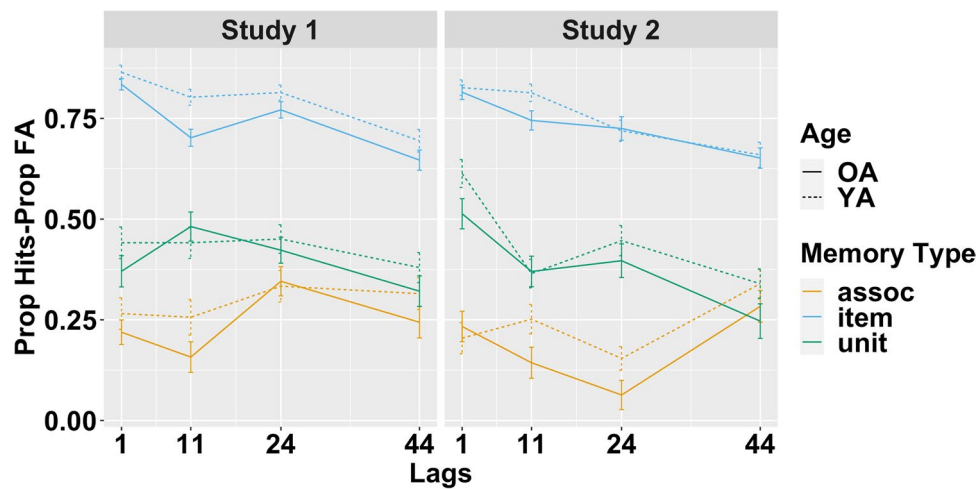
### Corrected recognition

The main analysis examined the effects of age, condition, and lag on corrected recognition (CR) using a 2 (between age: OA or YA)  $\times$  3 (within memory type: item, unitized,

associative)  $\times$  4 (within lag: 1, 11, 24, 44) mixed-model analysis of variance (ANOVA). Corrected recognition was calculated by subtracting each participant’s false-alarm rates from their hit rates in each condition, in each lag. There was a significant main effect of memory type on CR performance,  $F(1.74, 184.74) = 372.56$ ,  $p < .001$ ,  $\eta_p^2 = 0.78$ , such that items had higher CR compared with unitized pairs and associative pairs (all  $t$ s  $> 18.50$ , all  $p$ s  $< .001$ ). Unitized pairs also had higher CR compared with associative pairs,  $t(107) = 9.68$ ,  $p < .001$ . There was also a significant main effect of lag,  $F(3, 318) = 11.72$ ,  $p < .001$ ,  $\eta_p^2 = 0.10$ , such that the 1 lag had higher CR compared with the 44 lag,  $t(107) = 4.02$ ,  $p < .001$ . Additionally, the 24 lag had higher CR compared with the 11 lag and the 44 lag (all  $t$ s  $\geq 3.20$ , all  $p$ s  $< .05$ ). There was not a significant effect of age on corrected recognition,  $F(1, 106) = 3.32$ ,  $p = .07$ ,  $\eta_p^2 = 0.03$ .

There was a significant memory type by lag interaction,  $F(5.50, 583.02) = 9.32$ ,  $p < .001$ ,  $\eta_p^2 = 0.08$ . Pairwise  $t$  tests examining effects of memory type within lag show that in the 1 lag, items had higher CR compared with unitized and associative pairs (all  $t$ s  $> 15.23$ , all  $p$ s  $< .001$ ). In the 1 lag, the unitized pairs also had higher CR compared with associative trials,  $t(107) = 6.00$ ,  $p < .001$ . In the 11 lag, items had higher CR than unitized and associative trials (all  $t$ s  $> 9.77$ , all  $p$ s  $< .001$ ). Additionally, in the 11 lag, unitized pairs had higher CR compared with associative pairs,  $t(107) = 7.97$ ,  $p < .001$ . In the 24 lag, items had higher CR compared with unitized and associative trials (all  $t$ s  $> 12.54$ , all  $p$ s  $< .001$ ). In the 24 lag, unitized also had higher CR compared with the associative trials,  $t(107) = 3.45$ ,  $p = .002$ . In lag 44, items had higher CR compared with unitized and associative trials (all  $t$ s  $> 10.25$ , all  $p$ s  $< .001$ ). (For a breakdown of

<sup>2</sup> During analysis, we found two small errors in the hard coding of the task. One item from the 1 lag and one item from the 24 lag were labelled as lures when they were in fact targets. This was accounted for in the analyses and thus resulted in a slightly unbalanced number of targets in the item condition (i.e., 5 trials in the 1 and 24 lags and 7 in the 11 and 44 lags). This was fixed before collecting the second study.



**Fig. 2** The effect of memory type, age, and lag on corrected recognition. See results for relevant statistics. Assoc = associative; unit= unitization

results of hit and false-alarms rates, please see Supplemental Materials.)

Taken together, the above results support that unitized pairs have higher corrected recognition compared with associative pairs at all lags, apart from the 44 lag. Additionally, single items had higher corrected recognition compared with paired conditions at all lags. See Supplemental Material for corrected recognition broken down into false alarms and hits (Fig. 2).

### Experiment 1 discussion

Collapsed across age and lag, the results showed a main effect of memory type such that item memory had overall higher corrected recognition than unitized and associative memory. Additionally, unitized memory showed higher corrected recognition than associative memory. The difference between item and associative memory replicates a vast amount of past research in both young and older adults highlighting the differences in difficulty between the two types of memory (Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2004; Naveh-Benjamin et al., 2012; Old & Naveh-Benjamin, 2008). Most interestingly, is the fact that associative word pairs in the unitized condition showed higher corrected recognition compared with the nonunitized associative condition. Previous work in the field of unitization has suggested that the benefits of unitization in memory stem from its ability to represent associative pairs as a single item (Ahmad & Hockley, 2014; Ford et al., 2010; Quamme et al., 2007). Additionally, past work has shown that unitization can enhance associative memory in older adults' to be equitable to that of younger adult's performance (Ahmad et al., 2014).

Despite the overall benefit in memory for unitized word pairings compared with the nonunitized associative condition, we did not find that memory for unitized information

rose to the level seen for item memory. That is, items had overall higher corrected recognition at all lags due to unitized pairings having higher false-alarm rates (see Supplemental Material for false-alarm statistics). Taken together results suggest that unitization may operate as an effective means for engaging in associative binding without creating a single representation for the associative pairings (Parks & Yonelinas, 2015). Given that we did not find any differences in age groups, this conclusion may be especially critical to the field of aging, where age deficits in associative memory are typically prevalent compared with younger adults; however, further analysis is required to determine whether or not older and younger adults perform to the same degree (Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2012; Old & Naveh-Benjamin, 2008; Saverino et al., 2016). More specifically, it supports an interpretation of unitization that supported the concept as a "levels of unitization" continuum, which suggests that memory for unitized stimuli may fall on a spectrum between acting as an associative memory (two distinct items) or like item memory (like one cohesive item; Parks & Yonelinas, 2015). This may be critical to unitization literature such that if unitization operates as means for enhancing associative binding yet falling short of creating a single item representation in memory; somewhere between an item and an unrelated associative pair, it may engage hippocampal processes and not solely perirhinal cortex during encoding.

### Experiment 2

While results from Experiment 1 found that memory for the unitized condition was better than that of the associative condition at earlier lags, we did not see performance reach the level of item memory, as suggested by past literature



(Ahmad & Hockley, 2014; Bastin et al., 2013; Delhay & Bastin, 2016). One explanation for this finding may be that the word pairings were not viewed as a singular item (Parks & Yonelinas, 2015). While some work using unitized word pairs has physically separated the unitized word pairings (i.e., *black-bird*; Ahmad & Hockley, 2014), other work has used compound words (i.e., *mailbox*) as a form of unitization and found that they significantly outperformed unrelated associates (Giovanello et al., 2006). However, neither of these previous literatures had an item condition used as a point of comparison to see whether those compound words/pairings reached a similar performance to that of items or whether it just outperformed unrelated associates. In Experiment 2 we removed the spacing and dot between the two halves of the compound word in order to investigate whether the physical appearance of the word, presented as a compound word, would elevate memory in the unitized condition to that of item memory.

## Methods

### Participants

Participants provided a waiver of documentation of consent for a protocol approved by the Pennsylvania State University Institutional Review Board. Demographics were collected prior to beginning the task. Sixty-eight younger and 56 older adult participants were recruited and tested for Experiment 2. Younger adults were recruited from the psychology subject pool, where they participated in person at the Pennsylvania State University and received course credit for their effort. The older adults completed the tasks online through Prolific and were paid \$4.69 for participating. Subjects were removed from analyses on the grounds of no response rates over 10% of trials (10 young), memory at chance (four young; five old), bias to saying “old” (one young) or completing both versions of the task (one old). Thus, 53 younger adult participants ( $M_{\text{age}} = 19.86$ ,  $SD_{\text{age}} = 1.41$ ; 39 female;  $M_{\text{education}} = 13.64$ ,  $SD_{\text{education}} = 0.98$ ) and 50 older adult participants ( $M_{\text{age}} = 66.63$ ,  $SD_{\text{age}} = 4.76$ ; 34 female;  $M_{\text{education}} = 16.52$ ,  $SD_{\text{education}} = 2.25$ ) were included in the analyses. Younger adult participants identified as White (38), Black or African American (5), Asian (4), American Indian/Alaskan Native (1), or more than one race/other (5). Older adult participants identified as White (44), Black or African American (2), American Indian/Alaskan Native (1), or other/preferred not to answer (3). Both online and in-person testing followed the same protocol, with instructions printed on the screen. Following the task, participants were debriefed.

### Stimuli and design

The only difference in Experiment 2 was that the spacing and dot between the words in the unitized condition was

removed to create a compound word (see Fig. 1). Additionally, word order was rerandomized, such that words assigned to a given lag in Experiment 1 were assigned to a different random lag in Experiment 2. All other stimuli and design elements remained the same from Experiment 1 aside from being rearranged in the trial order.

### Procedure

All procedures were the same as in Experiment 1.

## Results

### Corrected recognition

The 2 (age: young and old)  $\times$  3 (memory: item, unitized, associative)  $\times$  4 (lag: 1, 11, 24, 44) mixed-model ANOVA examining corrected recognition revealed a significant main effect of memory type,  $F(2, 202) = 347.52$ ,  $p < .001$ ,  $\eta_p^2 = 0.78$ , such that items had higher CR compared with unitized and associative trials (all  $t_s > 16$ , all  $p_s < .001$ ). Additionally, unitized had higher CR compared with associative trials,  $t(102) = 11$ ,  $p < .001$ . There was also a significant main effect of lag,  $F(3, 303) = 21.57$ ,  $p < .001$ ,  $\eta_p^2 = 0.18$ , such that the 1 lag had higher CR compared with the other lags (all  $t_s > 5.31$ , all  $p_s < .001$ ). There was not a significant main effect of age group on CR rates,  $F(1, 101) = 3.91$ ,  $p = .051$ ,  $\eta_p^2 = 0.04$ .

There was also a significant memory type by lag interaction,  $F(5.4, 548) = 21.34$ ,  $p < .001$ ,  $\eta_p^2 = 0.17$ . Pairwise paired t-tests of memory type within lag showed that in the 1 lag, items had higher CR compared with unitized and associative trials (all  $t_s > 9.33$ , all  $p_s < .001$ ). In the 1 lag, unitized had higher CR compared with associative trials,  $t(102) = 11.39$ ,  $p < .001$ . In the 11 lag, items had higher CR compared with unitized and associative trials (all  $t_s > 12.92$ , all  $p_s < .001$ ). In the 11 lag, the unitized had higher CR compared with associative trials,  $t(102) = 5.13$ ,  $p < .001$ . In the 24 lag, items had higher CR compared with unitized and associative trials (all  $t_s > 8.65$ , all  $p_s < .001$ ). In the 24 lag, unitized had higher CR compared with associative trials,  $t(102) = 9.66$ ,  $p < .001$ . Finally, in the 44 lag, items had higher CR compared with unitized and associative trials (all  $t_s > 10.21$ , all  $p_s < .001$ ).

There was also a significant age by memory type by lag interaction,  $F(5.4, 548) = 2.19$ ,  $p = .004$ ,  $\eta_p^2 = 0.02$ . After running a series of two-way interactions, the three-way interaction was found to be driven mainly by interactions of memory type and lag within the age factor. In order to investigate these interactions further, we ran separate ANOVAs within each age group. Specifically, within each age group, a 3 (memory type)  $\times$  4 (lag) ANOVA was run. In younger adults (YAs), there was a main effect of memory type:  $F(2,$

104) = 191,  $p < .001$ ,  $\eta_p^2 = 0.79$ , such that items had higher CR compared with unitized and associative trials (all  $t_s \geq 13.8$ , all  $p_s < .001$ ). Additionally, unitized had higher CR compared with associative trials  $t(211) = 8.39$ ,  $p < .001$ . There was also a main effect of lag,  $F(3, 156) = 10$ ,  $p < .001$ ,  $\eta_p^2 = 0.16$ , such that the 1 lag had higher CR compared with the 11, 24, and 44 lags (all  $t_s \geq 2.99$ , all  $p_s \leq .019$ ).

There was also a significant memory type by lag interaction in younger adults,  $F(6, 312) = 13$ ,  $p < .001$ ,  $\eta_p^2 = 0.20$ . Pairwise  $t$  tests examining effects of memory type within lag, reveal that in the 1 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 5.40$ , all  $p_s < .001$ ). Unitized trials also had higher CR compared with associative trials,  $t(52) = 9.90$ ,  $p < .001$ . In the 11 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 10.76$ , all  $p_s < .001$ ). In the 24 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 5.56$ , all  $p_s < .001$ ). Additionally, in the 24 lag, unitized trials had higher CR compared with associative trials,  $t(52) = 6.99$ ,  $p < .001$ . In the 44 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 6.34$ , all  $p_s < .001$ ).

In older adults, there was a significant main effect of memory type,  $F(2, 98) = 160$ ,  $p < .001$ ,  $\eta_p^2 = 0.77$ , such that items had higher CR compared with unitized and associative trials (all  $t_s \geq 15.50$ , all  $p_s < .001$ ). Additionally, the unitized trials had higher CR compared with associative trials  $t(199) = 7.75$ ,  $p < .001$ . There was also a main effect of lag,  $F(3, 147) = 12$ ,  $p < .001$ ,  $\eta_p^2 = 0.19$ , such that the 1 lag had higher CR compared with the 11, 24, and 44 lags (all  $t_s \geq 4.77$ , all  $p_s < .001$ ).

In the older adults, there was also a significant memory type by lag interaction,  $F(5.1, 248) = 10$ ,  $p < .001$ ,  $\eta_p^2 = 0.17$ . Pairwise  $t$  tests examining effects of memory type within lag, reveal that in the 1 lag, items have higher CR compared with unitized and associative trials (all  $t_s \geq 8.07$ , all  $p_s < .001$ ). In the 1 lag, unitized had higher CR compared with associative trials,  $t(49) = 6.48$ ,  $p < .001$ . In the 11 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 7.71$ , all  $p_s < .001$ ). In the 11 lag, unitized trials had higher CR compared with associative trials,  $t(49) = 5.20$ ,  $p < .001$ . In the 24 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 6.69$ , all  $p_s < .001$ ). In the 11 lag, unitized trials had higher CR compared with associative trials,  $t(49) = 6.67$ ,  $p < .001$ . In the 44 lag, items had higher CR compared with unitized and associative trials (all  $t_s \geq 8.30$ , all  $p_s < .001$ ). For a breakdown of results of hit and false-alarm rates, please see Supplemental Materials.

Taken together, the above results support that unitized pairs have higher corrected recognition compared with associative pairs at all lags, apart from the 44 lag. Additionally, single items had higher corrected recognition compared with

paired conditions at all lags. See Supplemental Material for corrected recognition broken down into false alarms and hits. Fig. 2.

## Experiment 2 discussion

The only difference in the design across studies was the physical composition of the word pairings in the unitized condition. Specifically, the spacing between the unitized word pairings was removed in Experiment 2 such that the two words were arranged to form the single compound word that the two individual words created. This was done in an attempt to create greater cohesion amongst the word pairs and allow them to physically mirror a single word. Despite this design change, the results largely replicate Experiment 1. That is, across all lags, memory for items had higher corrected recognition compared with both the unitized and non-unitized associative conditions, with memory in the unitized condition also showing higher overall corrected recognition compared with nonunitized associative memory at earlier and middle lags (lags 1, 11, and 24). Taking a closer look at corrected recognition scores, we see that, like Experiment 1, differences across condition were driven by differences in false-alarm rates (see Supplemental Material for false-alarm statistics). Specifically, the unitized condition exhibited fewer false alarms at all lags compared with the nonunitized associative condition.

Like Experiment 1, the results provide evidence that unitized words boost associative memory in short term memory but fall to the level of more traditional nonunitized associative memory with time and interference. The results also suggest that regardless of how similar two items pre-experimentally unitized are, they do not function as a single representation in memory. That is, when they are recombined to form a lure, memory for the original unitized pairing is not strong enough to reject the new lure pairing at the same correct rejection rate as that of novel words. Thus, we conclude that regardless of physical similarity to an item, memory discrimination for unitized compound words may never quite reach the same performance of items. As such, the results further support the idea of a continuum of unitization (Parks & Yonelinas, 2015), such that unitization may fall somewhere between that of items and associations.

## General discussion

The current set of studies was designed to investigate how memory for unitized pairings functions across time and interference as we age, in comparison to memory for nonunitized associations and items. Results were largely consistent across both studies, showing that corrected recognition for both words and preexperimentally unitized word pairings

was overall better than memory for nonunitized word pairs. Yet, despite this advantage of unitization compared with nonunitized associative memory, memory in the unitized condition fell below that of item memory. These differences in corrected recognition rates were largely due to differences in false alarm rates across conditions, where false alarms for rearranged yet novel word pairings were higher in the nonunitized associative compared with the unitized condition and lowest in the item condition (see Supplemental Materials for false alarm statistics). Taken together the results indicate that there is a benefit to preexperimentally unitized item pairings over and above that of nonunitized associative pairings within a certain time range (see section on role of interference for more regarding time effects); yet unitization does not operate by creating a single item from the associative pairings. Additionally, the current work indicates that both young and especially older adults benefit from unitization, such that their memory performance for unitized pairs is over and above that of associative pairs across increasing time and interference. While there were no significant differences between young and older adults, further work is necessary to confirm that younger and older adults benefit from unitization to the same degree.

Past work has suggested that unitization enhances associative memory by allowing unrelated item pairings to be represented in memory at the level of a single item (Ahmad & Hockley, 2014; Bastin et al., 2013; Delhay & Bastin, 2016). Even the very earliest of work in the field of unitization suggests that this benefit may arise from related words utilizing less space in short-term memory compared with unrelated words (Fritzen, 1974). Both areas of unitization research rest on the idea that larger units are chunked or represented as smaller units in memory stores. This conclusion is supported by neuroimaging work that finds that memory for nonunitized associative pairs is typically localized to the hippocampal regions while memory for items and unitized associative pairs tends to be processed in perirhinal cortex (Staresina & Davachi, 2008, 2009, 2010). However, no experiment, to our knowledge, has directly compared memory for unitized pairings with memory for items in order to make this conclusion. Additionally, implicit in the foregoing explanation is the notion that a single unit cannot be broken apart, and hence unitized information remains inherently bound in memory. The current results suggest that this is not how unitized information is stored. That is, if preexperimentally unitized word pairings, such as *mailman* and *shoebox* were stored as single representations, then the rearranged word of *mailbox* should not incur high false-alarm rates but be viewed as a novel lure during retrieval. The fact that the rearranged lure, *mailbox* is often false alarmed to suggests that unitized pairings may be stored as separate, yet bound items in memory.

Thus, while it may be that unitized pairings are not bound into a single item representation, this evidence suggests that they may be bound better than nonunitized associative word pairings, as they are less likely to be misidentified once broken apart and rearranged. Furthermore, the results suggest that both younger and older adults are able to take advantage of unitization to improve associative memory, particularly at earlier lags. This finding is consistent with a view of unitization that suggests unitization operates by creating higher cohesion between the items in a pairing compared with associative pairs (Ahmad & Hockley, 2014; Bastin et al., 2013; Delhay & Bastin, 2016; Diana et al., 2008; Quamme et al., 2007). This stronger bond may allow for participants to correctly reject a recombined unitized lure more often than is found with unrelated, associative pairings. This past work has also suggested that the use of unitization is especially beneficial to older adults, allowing them to make use of item memory processes (Ahmad et al., 2014; Delhay & Bastin, 2016). While the current results cannot speak directly to the mechanism underlying improvements in the unitized condition, or whether this is the same across age groups, the evidence suggests that unitization reflects a difference in processing compared with traditional associative memory tasks, and one that is not reflective of item memory. This insight is also consistent with the interpretation that unitization operates as a continuum in memory, allowing for enhanced associative binding, but not distilling down item–item pairs to a single representation (Parks & Yonelinas, 2015). To explore this idea further, future work should manipulate the degree of relatedness across the unitized pair as well as the manner in which the stimuli are presented together (i.e., physical distance, location, etc.).

Interestingly, most of the prior work in the realm of unitization (Ahmad & Hockley, 2014; Bastin et al., 2013; Delhay & Bastin, 2018; Diana et al., 2008; Graf & Schacter, 1989) typically finds higher hit rates driving performance benefits to unitization, while false alarms remain relatively similar between unitized and associative pairings. The current results saw significant differences in false-alarm rates across conditions, with false alarms being reduced in the unitized compared with the associative condition (see Supplemental findings). Differences between the current results and past work may be due in part to how participants were tested. Some of the foregoing studies do not use recombined lures, but rather implement a cued recall task for retrieval (Bastin et al., 2013; Graf & Schacter, 1989). As such, the cued recall retrieval task in previous work relied less on recognition than our task. Thus, the differences in hits rather than false-alarm rates found in some previous work may have been due to hippocampal-based retrieval processes rather than familiarity and perirhinal cortex-based recognition processes. The current task equated that here and allowed for recognition and familiarity-based retrieval



processes to be utilized to influence both false-alarm and hit rates. However, this is not consistent across all the literature, thus future work in understanding the benefits of unitization should also explore the effect of testing format.

While the current results support the finding that unitization enhances associative memory above that of nonunitized associative memory, there is also evidence that this benefit is limited to shorter term memory. That is, the results in both Studies 1 and 2 suggest that memory in the unitized condition is better than the associative condition at earlier (1, 11, 24), but not later (44) lags, indicating that time and interference interact with condition to diminish the unitization benefit to corrected recognition. Prior work using this paradigm utilized these lags to represent shorter term memory and shorter long-term memory processes (Kuhlmann et al., 2021). Thus, the current results suggest that unitization provides more support to associative memory with respect to shorter rather than longer term memory processes, with this benefit eventually subsiding once enough time and interference from additional information has been encountered.

With respect to the role of interference, as measured by lags, the current results largely replicates the previous work done by Kuhlmann et al. (2021). Specifically, Kuhlmann and colleagues found that while item memory showed a linear decline with lag, associative memory showed stability at earlier lags, and a decline with less stability in corrected recognition only at later lags. In the current results, the item conditions exhibited a linear decline with increasing time and interference, while the associative condition shows less stable trends across time and interference similar to what Kuhlmann found. This same linear decline was also observed in the unitized condition, specifically in study 2 where the unitized words were presented as a single compound word. Taken together, this may suggest that, while unitized words do not reach the overall level of performance as that of items, they may function in a manner more similar to that of an item, with respect to the effect of interference and time. Furthermore, the slightly different associative memory patterns in later lags may have been a difference in how items and associative pairings were presented and tested. In the Kuhlmann experiment, item memory was tested as a component of associative memory, whereby item targets were always formerly one half of an associative pair (Kuhlmann et al., 2021). In the current experiment, we utilized single items, associative, and unitized pairings across both study and test trials, thus leading to a larger number of trials. Future work should attempt to replicate previous findings more directly to assess the effect of interference on associative memory and determine how different testing parameters affect memory performance across conditions (i.e., item, associative, and unitized memory pairs). The weaker stability in the later lags could also be attributed to sampling variance and may stabilize with more participants,

or with additional trials. Future work could aim to replicate this work with a larger number of participants and trials to determine whether this fully replicates previous work.

### Limitations and future directions

There were a few limitations to the current experiment. Item lures were foils rather than related or recombined lures like the unitized and associative conditions, which may be a confounding factor and what is driving the high rates of overall item corrected recognition. Future work could potentially create a related lure for items in order to get a better understanding of how false alarms impact item corrected recognition. Additionally, with respect to the stimuli used in the study, we did not directly consider word concreteness or imageability when developing the task largely due to the constraints placed on the unitized condition and need to use compound words. Future work could attempt to develop word lists that match in terms of concreteness and imageability. Additionally, with respect to the unitized words, we drew our stimuli from previously published work that also used compound word pairs. As such, these word pairs are, for experimental purposes, considered to be preexperimentally unitized (Ahmad et al., 2014; Ahmad & Hockley, 2014; Delhay & Bastin, 2018; Quamme et al., 2007). Thus, we did not include any form of encoding instruction for the unitized condition that differed from the other two memory conditions, as one of our goals was to examine associative memory in this manner, and using differing instructions would be counter to the current study's questions. However, future work could induce unitization via instructions to determine whether participants are able to unitize spontaneously.

Additionally, with respect to task design, we did not measure familiarity directly, which is often referenced in unitization literature. Future work could utilize a remember-know-new response paradigm to examine familiarity more directly. In addition to an RKN paradigm, the effect of lag in memory across the three trial types could be examined using a free recall test. While we used a continuous recognition paradigm in order to examine the effects of time and interference on the different memory types, recall tests might lead to different conclusions regarding discriminability across conditions, given that false alarms are less common in free recall tests. Future work could also examine these memory types with larger time intervals. Here, our 24 lag was equivalent to about 2 minutes, and our 44 lag was equivalent to about 4 minutes between study and test trials. Thus, the current experiment may not encompass how these types of memory may act across days or even hours.

Finally, older and younger adults were tested in different settings, with younger adults being tested in person and older adults tested online. Whether this contributed to the absence of differences, it is hard to say. It may be that older

adults who were able to navigate online testing protocols may have higher levels of cognitive abilities compared with the general population of older adults that typically volunteer from the community, thus allowing them to perform similarly to younger adults. Future work should replicate the current findings by testing a community-dwelling older adult sample that may be more representative of previous studies.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.3758/s13421-022-01389-w>.

**Acknowledgements** We thank Aislinn Thomas, Joseph Drexler, and Sophia Patronelli for help with data collection and analyses. This work was supported by the National Institute of Aging under grant 1R56AG070014-01A1 awarded to N.A.D.

**Data availability** The datasets generated and analyzed during the current experiments are available from the corresponding author on reasonable request. None of the experiments was preregistered.

**Code availability** Code is available upon request from the corresponding author.

## Declarations

**Conflicts of interest** No potential conflict of interest was reported by either author.

## References

- Ahmad, F. N., & Hockley, W. E. (2014). The role of familiarity in associative recognition of unitized compound word pairs. *Quarterly Journal of Experimental Psychology*, 67(12), 2301–2324. <https://doi.org/10.1080/17470218.2014.923007>
- Ahmad, F. N., Fernandes, M., & Hockley, W. E. (2014). Improving associative memory in older adults with unitization. *Aging, Neuropsychology, and Cognition*. <https://doi.org/10.1080/13825585.2014.980216>
- Ashford, J. W., Gere, E., & Bayley, P. J. (2011). Measuring memory in large group settings using a continuous recognition test. *Journal of Alzheimer's Disease*, 27(4), 885–895. <https://doi.org/10.3233/JAD-2011-110950>
- Bader, R., Mecklinger, A., Hopstädter, M., & Meyer, P. (2010). Recognition memory for one-trial-unitized word pairs: Evidence from event-related potentials. *NeuroImage*, 50(2), 772–781. <https://doi.org/10.1016/j.neuroimage.2009.12.100>
- Bastin, C., Diana, R. A., Simon, J., Collette, F., Yonelinas, A. P., & Salmon, E. (2013). Associative memory in aging: The effect of unitization on source memory. *Psychology and Aging*, 28(1), 275–283. <https://doi.org/10.1037/a0031566>
- Berman, S., Friedman, D., & Cramer, M. (1991). ERPs during continuous recognition memory for words and pictures. *Bulletin of the Psychonomic Society*, 29(2), 113–116. <https://doi.org/10.3758/BF03335209>
- Chen, T., & Naveh-Benjamin, M. (2012). Assessing the associative deficit of older adults in long-term and short-term/working memory. *Psychology and Aging*, 27(3), 666–682. <https://doi.org/10.1037/a0026943>
- Delhay, E., & Bastin, C. (2016). The impact of aging on associative memory for preexisting unitized associations. *Aging, Neuropsychology, and Cognition*. <https://www.tandfonline.com/doi/full/10.1080/13825585.2016.1263725>
- Delhay, E., & Bastin, C. (2018). The impact of aging on associative memory for preexisting unitized associations. *Aging, Neuropsychology, and Cognition*, 25(1), 70–98. <https://doi.org/10.1080/13825585.2016.1263725>
- Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 730–740. <https://doi.org/10.1037/0278-7393.34.4.730>
- Ford, J. H., Verfaellie, M., & Giovanello, K. S. (2010). Neural correlates of familiarity-based associative retrieval. *Neuropsychologia*, 48(10), 3019–3025. <https://doi.org/10.1016/j.neuropsychologia.2010.06.010>
- Fritzen, J. (1974). Unitization of related words in short-term store. *Journal of Experimental Psychology*, 103(4), 638–645. <https://doi.org/10.1037/h0037195>
- Giovanello, K. S., Keane, M. M., & Verfaellie, M. (2006). The contribution of familiarity to associative memory in amnesia. *Neuropsychologia*, 44(10), 1859–1865. <https://doi.org/10.1016/j.neuropsychologia.2006.03.004>
- Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 930–940. <https://doi.org/10.1037/0278-7393.15.5.930>
- Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: The role of active forgetting in memory. *Trends in Cognitive Sciences*, 17(3), 111–120. <https://doi.org/10.1016/j.tics.2013.01.001>
- Hennessey, J. P., Knowlton, B. J., & Castel, A. D. (2018). The effects of value on context-item associative memory in younger and older adults. *Psychology and Aging*, 33(1), 46–56. <https://doi.org/10.1037/pag0000202>
- Hockley, W. E. (1982). Retrieval processes in continuous recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(6), 497–512. <https://doi.org/10.1037/0278-7393.8.6.497>
- Jones, T. C., & Atchley, P. (2002). Conjunction error rates on a continuous recognition memory test: Little evidence for recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(2), 374–379. <https://doi.org/10.1037/0278-7393.28.2.374>
- Kim, M.-S., Kim, J.-J., & Kwon, J. S. (2001). The effect of immediate and delayed word repetition on event-related potential in a continuous recognition task. *Cognitive Brain Research*, 11(3), 387–396. [https://doi.org/10.1016/S0926-6410\(01\)00011-8](https://doi.org/10.1016/S0926-6410(01)00011-8)
- Kuhlmann, B. G., Brubaker, M. S., Pfeiffer, T., & Naveh-Benjamin, M. (2021). Longer resistance of associative versus item memory to interference-based forgetting, even in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(3), 422–438. <https://doi.org/10.1037/xlm0000963>
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1170–1187. <https://doi.org/10.1037/0278-7393.26.5.1170>
- Naveh-Benjamin, M., Guez, J., Kilb, A., & Reedy, S. (2004). The associative memory deficit of older adults: Further support using face-name associations. *Psychology and Aging*, 19(3), 541–546. <https://doi.org/10.1037/0882-7974.19.3.541>
- Naveh-Benjamin, M., Maddox, G. B., Jones, P., Old, S., & Kilb, A. (2012). The effects of emotional arousal and gender on the associative memory deficit of older adults. *Memory & Cognition*, 40(4), 551–566. <https://doi.org/10.3758/s13421-011-0169-x>
- Old, S. R., & Naveh-Benjamin, M. (2008). Memory for people and their actions: Further evidence for an age-related associative deficit. *Psychology and Aging*, 23(2), 467–472. <https://doi.org/10.1037/0882-7974.23.2.467>
- Parks, C. M., & Yonelinas, A. P. (2015). The importance of unitization for familiarity-based learning. *Journal of Experimental*

- Psychology. Learning, Memory, and Cognition*, 41(3), 881–903. <https://doi.org/10.1037/xlm0000068>
- Poon, L. W., & Fozard, J. L. (1980). Age and word frequency effects in continuous recognition memory. *Journal of Gerontology*, 35(1), 77–86. <https://doi.org/10.1093/geronj/35.1.77>
- Portrat, S., Barrouillet, P., & Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1561–1564. <https://doi.org/10.1037/a0013356>
- Quamme, J. R., Yonelinas, A. P., & Norman, K. A. (2007). Effect of unitization on associative recognition in amnesia. *Hippocampus*, 17(3), 192–200. <https://doi.org/10.1002/hipo.20257>
- Roediger, H. L. (1985). Remembering Ebbinghaus. *Contemporary Psychology: A Journal of Reviews*, 30(7), 519–523. <https://doi.org/10.1037/023895>
- Saverino, C., Fatima, Z., Sarraf, S., Oder, A., Strother, S. C., & Grady, C. L. (2016). The associative memory deficit in aging is related to reduced selectivity of brain activity during encoding. *Journal of Cognitive Neuroscience*, 28(9), 1331–1344. [https://doi.org/10.1162/jocn\\_a\\_00970](https://doi.org/10.1162/jocn_a_00970)
- Staresina, B. P., & Davachi, L. (2008). Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. *Journal of Cognitive Neuroscience*, 20(8), 1478–1489. <https://doi.org/10.1162/jocn.2008.20104>
- Staresina, B. P., & Davachi, L. (2009). Mind the gap: Binding experiences across space and time in the human hippocampus. *Neuron*, 63(2), 267–276. <https://doi.org/10.1016/j.neuron.2009.06.024>
- Staresina, B. P., & Davachi, L. (2010). Object unitization and associative memory formation are supported by distinct brain regions. *Journal of Neuroscience*, 30(29), 9890–9897. <https://doi.org/10.1523/JNEUROSCI.0826-10.2010>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.