



# Individual differences in working memory capacity and search efficiency

Ashley L. Miller<sup>1</sup> · Nash Unsworth<sup>1</sup>

Published online: 29 May 2018  
© Psychonomic Society, Inc. 2018

## Abstract

In two experiments, we examined how various learning conditions impact the relation between working memory capacity (WMC) and memory search abilities. Experiment 1 employed a delayed free recall task with semantically related words to induce the buildup of proactive interference (PI) and revealed that the buildup of PI differentially impacted recall accuracy and recall latency for low-WMC and high-WMC individuals. Namely, the buildup of PI impaired recall accuracy and slowed recall latency for low-WMC individuals to a greater extent than what was observed for high-WMC individuals. To provide a circumstance in which previously learned information remains relevant over the course of learning, Experiment 2 required participants to complete a multitrial delayed free recall task with unrelated words. Results revealed that with increased practice with the same word list, WMC-related differences were eventually eliminated in interresponse times (IRTs) and recall accuracy, but not recall latency. Thus, despite still accumulating larger search sets, low-WMC individuals searched LTM as efficiently as high-WMC individuals. Collectively, these results are consistent with the notion that under normal free recall conditions, low-WMC individuals search LTM less efficiently than do high-WMC individuals because of their reliance on noisy temporal–contextual cues at retrieval. However, it appears that under conditions in which previously learned items remain relevant at recall, this tendency to rely on vague self-generated retrieval cues can actually facilitate the ability to accurately and quickly recall information.

**Keywords** individual differences · memory · recall · working memory

The ability to accurately retrieve specific experiences from our past, referred to as episodic memories (Tulving, 2002), is essential for a variety of everyday tasks. Oftentimes, though, merely retrieving accurate information is not enough. We are frequently burdened with the additional need to recall this information in a timely manner. For instance, we commonly find ourselves in social situations in which we try to impress colleagues with witty or intelligent conversation, or in generating ideas at work. Given the inescapable need to accurately and quickly recall items from long-term memory (LTM), it is unsurprising that memory research has long been interested in better understanding search-efficiency processes (Raaijmakers & Shiffrin, 1980; Shiffrin, 1970). More recent work, however, reveals that substantial variation exists in the ability to successfully and quickly locate sought-after information in LTM (Unsworth, 2007; Unsworth & Engle, 2007; Unsworth,

Spillers, & Brewer, 2012). As a result, the present investigation examined whether we could first exaggerate, then eliminate, these differences in memory-search abilities among high-ability (e.g., people with high working memory capacity) and low-ability (e.g., people with low working memory capacity) individuals. In doing so, we hoped to better elucidate the reasons for which some individuals struggle to normally search the contents of LTM.

One of the oldest and most popular paradigms used to examine episodic memory performance is free recall (Tulving, 1968). Free recall refers to circumstances in which individuals are presented with information and later asked to recall that information in no particular order and under conditions providing minimal environmental support (i.e., no external cues are given to guide retrieval). Free recall tasks, therefore, demand substantial self-initiated, controlled processing at retrieval (see Craik, 1986; Salthouse, 2001). Of interest to the present study, much of our current understanding pertaining to efficient memory search processes has been driven by search models created in an attempt to better clarify the nature of free recall. Notably, many of these models incorporate the use of self-generated retrieval cues (Davelaar & Raaijmakers, 2012; Raaijmakers & Shiffrin, 1980; Rohrer &

---

✉ Ashley L. Miller  
amiller8@uoregon.edu

<sup>1</sup> Department of Psychology, 1227 University of Oregon,  
Eugene, OR 97403, USA

Wixted, 1994; Wixted & Rohrer, 1994), which ease the process of locating desired items in LTM by confining the search to a subset of activated memory representations associated with the cue. This subset of activated information determined by the retrieval cue is commonly referred to as the search set, and search models of free recall assume that during recall, memory representations must be sampled from the search set based on a relative strength rule (Raaijmakers & Shiffrin, 1980; Shiffrin, 1970). That is, each item within the search set competes to be sampled in such a way that a single item's chances of being sampled depends not only on the strength of that specific item but also on the strength of the other items in the search set. Recovery, on the other hand, occurs once the activation of a sampled item exceeds some critical threshold, meaning sampled items whose strengths do not exceed the threshold cannot be recovered into consciousness and are subsequently never retrieved (known as the absolute strength rule; Rohrer, 1996).

One means of examining the breadth of search, or the number of items being sampled, is by assessing an individual's recall latency. Recall latency refers to the time point in a recall period when a successful response is produced; hence, mean recall latency is the average time taken to recall items. Note that the random search model assumes that sampling of memory representations within the search set operates at a constant rate, with replacement (Wixted & Rohrer, 1994). So after an item has been sampled, the chances of that particular item being selected on the next sample are unchanged because the item reassumes its place within the search set and may consequently be sampled again. Given this assumption, the random search model predicts that larger search sets should be associated with longer recall latencies (Wixted & Rohrer, 1994). Indeed, this seems to be the case; as the length of a list increases, so does mean recall latency (Rohrer & Wixted, 1994). Presumably, longer lists result in more target items being included in the search set, thereby increasing the amount of time needed to sample each given item. Similarly, Wixted and Rohrer (1993) revealed that as proactive interference increased across lists, recall latency also increased. More time was needed to sample items when the search set was artificially inflated by the presence of intrusions. Collectively, these findings suggest that recall latency is theoretically associated with how people search for items in LTM by indexing the relative strength of all to be sampled items within the search set.

An additional variable of interest commonly related to recall latency is interresponse times (IRTs), defined as the time between consecutive retrievals. Research (Murdock & Okada, 1970; Rohrer, 1996; Rohrer & Wixted, 1994) demonstrates that IRTs are determined by the number of recoverable items in one's search set that are not yet recalled, meaning IRTs grow exponentially as a function of output position

(Murdock & Okada, 1970; Rohrer & Wixted, 1994; Wixted & Rohrer, 1993, 1994). Assuming items are constantly sampled with replacement, it becomes more likely that each subsequent sample will result in recovering an already sampled item (i.e., repetition) or an intrusion. If subsequent samples result in recovering a series of intrusions, IRTs would be expected to increase. In support of this claim, Unsworth (2016) demonstrated that a more focused search set containing less irrelevant information was associated with faster mean IRTs when compared with a larger search set made up of more intrusions. Therefore, IRTs provide additional unique information about the efficiency of one's memory search by providing perhaps the most accurate portrayal of the time course of retrieval.

Taken altogether, it is apparent that the time taken to recall items, whether it be indexed via recall latencies or IRTs, is related to how an individual searches the contents of LTM. Of relevance to the present study, examining recall latencies and IRTs is especially useful when trying to determine the efficiency of one's memory search, with a more efficient search associated with faster recall latencies and IRTs in conjunction with recall of more target items and fewer errors (i.e., repetitions and intrusions). However, there exists considerable variation in these abilities, particularly in terms of the breadth of search. As a result, the present study was largely concerned with further examining one individual differences variable related to search-efficiency processes: working memory capacity (WMC; Unsworth, 2007; Unsworth & Engle, 2007; Unsworth et al., 2012).

Prior research has consistently shown that when a cue-dependent search through LTM is undertaken, high-WMC individuals display better LTM performance than do low-WMC individuals (Kane & Engle, 2000; Rosen & Engle, 1997). For example, Unsworth (2007, 2009a, 2009b) has revealed that low-WMC individuals consistently display worse recall accuracy than do high-WMC individuals. Critically, in each case, the inferior recall performance of low-WMC individuals is also characterized by slower recall latencies and IRTs. To further illustrate this point, we analyzed a prior unpublished data set from our lab containing 164 participants recruited from the University of Georgia's human subject pool. All participants completed three complex span tasks—the operation span task (Ospan), the symmetry span task (Symspan), and the reading span task (Rspan)—followed by a delayed free recall task consisting of six word lists each composed of 10 unrelated words. Consistent with prior work (Unsworth, 2007, 2009a, 2009b; Unsworth & Engle, 2007; Unsworth et al., 2012), higher WMC was associated with recall of more target items ( $r = .22, p < .01$ ) and less errors ( $r$  with repetitions =  $-.26, p < .01$ ;  $r$  with intrusions =  $-.20, p < .05$ ), as well as faster recall latencies ( $r = -.23, p < .01$ ) and IRTs ( $r = -.26, p < .01$ ). This pattern of results goes against the notion that individuals take longer to recall items when

more correct items are located within their search sets (i.e., it takes longer to recall items because they have more items to output). If this were the case, mean recall latency should be significantly positively correlated with recall accuracy, as well as WMC. The positive correlation with WMC would be expected due to the finding that higher WMC is related with superior recall of more target items. Therefore, the results reported above suggest that under normal free recall conditions, low-WMC individuals search through larger search sets due to the inclusion of more irrelevant items. In contrast, high-WMC individuals appear to search through smaller search sets, reflecting an improved ability to focus the search to target items in LTM.

Given the notion that WMC is associated with differences in the size of one's search set (Unsworth, 2007; Unsworth & Engle, 2007; Unsworth et al., 2012), the present study sought to examine whether we could modify this effect. Namely, Experiment 1 used a traditional proactive interference (PI) buildup task to assess whether inducing PI could intensify low-WMC individuals' tendency to search through more irrelevant items at retrieval. If low-WMC individuals are less able to constrain the search to target list items in LTM than are high-WMC individuals, the buildup of PI should interfere with recall latency and recall accuracy more for these individuals. Experiment 2 examined whether changing the relevance of previously learned material could actually decrease WMC-related differences in other aspects of the memory search with multitrial free recall. If low-WMC individuals typically incorporate more previous-list items in their search sets than do high-WMC individuals, conditions of multitrial free recall should still result in low-WMC individuals searching through larger search sets. But, rather than searching through more irrelevant items, low-WMC individuals may instead search through more target item copies (see Bower, 1967; Hintsman, 1988; Unsworth, 2015). If this is the case, we should observe a larger improvement in recall accuracy and faster IRTs across trials for low-WMC individuals than high-WMC individuals. Understanding how various learning conditions differentially impact dynamics of the memory search for high WMC and low-WMC individuals will allow us to better discern why some individuals struggle to normally engage in an efficient memory search.

## Experiment 1

The purpose of Experiment 1 was to examine how the buildup of PI—the phenomena by which previously learned material interferes with our ability to learn new material (Underwood, 1957)—influences WMC-related differences in search set size. Recall that retrieval cues are one means by which individuals delimit the search to relevant items in LTM (Shiffrin, 1970; Shiffrin & Atkinson, 1969). These cues may take on

many forms, such as contextual, categorical, or temporal cues. But, in the case of episodic memory tasks, temporal-contextual cues prove to be most advantageous. For instance, in traditional delayed free-recall tasks, participants must remember a series of unrelated word lists, with each list associated with its own contextual features. So, for optimal recall performance on these tasks, individuals must be able to properly distinguish current-list items from previous-list items via the use of appropriately targeted temporal-contextual retrieval cues. Under conditions of PI, though, it becomes progressively more difficult to correctly discriminate current-list items from previous-list items. That is, as PI builds across a task, participants are less able to constrain the search to the current list because they are incorporating previously presented material into their search sets, leading to impaired recall accuracy and slowed retrieval (see Wixted & Rohrer, 1993).

Prior research has further revealed that WMC-related differences in search set size are due to differences in the effectiveness of one's temporal—contextual cues at retrieval (see Unsworth, 2007; Unsworth & Engle, 2007; Unsworth et al., 2012). Given this notion, the buildup of PI should amplify low-WMC individuals' inability to correctly discriminate current-list items from previous-list items. Indeed, some prior work is consistent with this notion (see Kane & Engle, 2000; Unsworth, 2010a). Other work (Kliegl, Pastotter, & Bauml, 2015), however, has only partially supported this view. Specifically, Kane and Engle (2000) and Unsworth (2010a) revealed larger PI effects in recall accuracy for low-WMC individuals than high-WMC individuals, whereas Kliegl et al. (2015) observed no WMC differences in the PI effect in recall accuracy. Instead, WMC only moderated the PI effect in recall latencies. Since (1) only one study has examined WMC-related differences in the PI effect on recall latency and (2) prior research has yielded conflicting results pertaining to WMC-related differences in the PI effect on recall accuracy, we sought to better assess whether the buildup of PI magnifies WMC-related differences in both of these aspects of the memory search.

## Method

### Participants and procedure

A total of 161 participants (61% female) were recruited from the human subject pool at the University of Georgia. All participants were between the ages of 18 and 35 years ( $M = 19.20$  years,  $SD = 1.71$ ). Nonnative English speakers were excluded from the study, and all eligible participants were awarded course research credit for participation. After obtaining informed consent and demographic information, all participants completed three complex span tasks: the operation span task (Ospan), the symmetry span task (Symspan), and the reading span task (Rspan). Upon completion of the WMC tasks,

participants were then administered a delayed free recall task consisting of six word lists each composed of 10 semantically related words to induce the buildup of PI. Of note, participants completed the tasks reported herein as part of a larger experimental test battery lasting approximately 2 hours. Since the other tasks administered during the experimental session do not relate to the current study, they are not reported (see Unsworth, 2010b).

### WMC tasks

**Ospan** Participants solved a series of elementary math problems while remembering unrelated letters. First, on a computer, participants were presented with a math operation—for example,  $(4 \times 1) + 2 = ?$ —in which they had to click the mouse to indicate that they had solved the problem. A new screen then appeared with an answer to the math solution (e.g., 6), whereby participants had to indicate if the answer listed onscreen was correct or incorrect via mouse click (e.g., in the case above, the answer 6 would be correct). Upon completion of the math operation, participants were then presented with a letter (e.g., F, H, J, K, L, N, P, Q, R, S, T, and Y) for 1 second. Immediately following letter presentation, the next math problem was presented. Set sizes varied randomly from three to seven math operation/letter strings, and participants had to complete three trials of each set size for a total possible score of 75. At recall for each set, letters from the corresponding set had to be recalled in order by selecting the relevant letters (see Unsworth, Heitz, Schrock, & Engle, 2005, for more details).

**Symspan** Participants solved symmetry judgements while remembering the location of a sequence of red squares within a matrix. Symmetry judgements consisted of an  $8 \times 8$  matrix of squares in which some of the squares were filled black and the remaining squares remained white. Participants indicated whether the pattern created by the filled squares was symmetrical about the vertical axis. Once participants indicated whether they believed the pattern was symmetrical or nonsymmetrical, participants were shown a  $4 \times 4$  matrix with one of the cells filled red for 650 ms. Immediately following the presentation of the red square matrix, the next symmetry judgement trial began. Set sizes randomly ranged from two to five, and there were three trials of each set size for a total possible score of 42. Participants were asked to recall the sequence of red-square locations based on the order in which they were presented across the corresponding trial. Participants indicated the appropriate location of each red square by clicking on cells in an empty matrix (see Unsworth, Redick, Heitz, Broadway, & Engle, 2009, for more details).

**Rspan** While remembering the same unrelated letters as in the Ospan, participants provided judgements about a series of sentences. More specifically, participants read a sentence

containing 10 to 15 words and determined whether or not the sentence made sense to them (e.g., “Every now and then I catch myself swimming blankly at the wall”). Nonsense sentences were created by modifying a single word from an otherwise ordinary sentence (e.g., changing “staring” to “swimming” in the case above). Upon indicating whether the sentence made sense or not, participants were then presented with a letter for 1 second. Set sizes randomly varied from three to seven sentence/letter strings, and participants had to complete three trials of each set size for a total possible score of 75. At recall for each set, letters from the corresponding set had to be recalled in order by selecting the appropriate letters (see Unsworth et al., 2009, for more details).

**Composite WMC score** The following analyses involving WMC used a WMC composite score created for each participant by averaging participants’ standardized *Z* scores on each of the complex span tasks. This variable allowed us to treat WMC as a continuous variable in all analyses. For graphical purposes only, WMC was utilized as a categorical variable by using a quartile split in which the top 25% of performers on the WMC tasks were classified as high-WMC individuals, and the lowest 25% performers were considered low-WMC individuals.<sup>1</sup>

### Delayed free recall semantically related words

Participants were administered six lists of 10 semantically related words, each broken down into two blocks. One block consisted of three word lists composed of nouns from the semantic category fruits, whereas the other block contained three word lists made up of nouns from the semantic category professions (see Kane & Engle, 2000). The first three lists allowed for the buildup of PI, and the first list in the ensuing block (i.e., List 4) allowed for a release from PI. All words were presented onscreen for 1 second each. Following presentation of the last word within each list, a 15-second distractor task began that required participants to count backwards by 3s as quickly and accurately as possible from a three-digit number presented onscreen. Participants were further instructed to write the numbers down as they proceeded through the distractor task. After the distractor task, three question marks appeared at the center of the screen to prompt participants to recall as many words as possible within a 45-second window. Participants were instructed to type as many words as they could remember from the corresponding list in any order they wished. RTs were measured with respect to when participants pressed enter after each word was typed (see also Mickes, Seale-Carlisle, & Wixted, 2013). A participant’s score was

<sup>1</sup> Using a criterion of 80% accuracy on the processing component across the three span tasks similar to prior research (Unsworth, 2007; Unsworth & Engle, 2008) suggested that no participants had to be excluded in either experiment.



the total number of items recalled correctly, and mean recall latency, total repetitions, and total intrusions were likewise recorded.

## Results

Consistent with prior research (Unsworth, 2007, 2009a, 2009b; Unsworth & Engle, 2007; Unsworth et al., 2012), higher WMC was associated with overall better recall accuracy ( $r = .16, p = .05$ ), faster recall latencies ( $r = -.25, p < .01$ ), and faster IRTs ( $r = -.17, p < .05$ ). Descriptive statistics and estimates of reliability for all tasks are listed in Table 1.<sup>2</sup> For correlations between all measures, see Table 2.

### Proportion correct recall

First, a repeated-measures ANCOVA was used to examine proportion of correctly recalled items as a function of List Type (within-subjects factor) and WMC (covariate). Note that there were six total lists used in the delayed free recall task (see Fig. 1a). Because the lists were broken into two blocks, with List 1 of each block representing minimum buildup of PI and List 3 within each block representing maximum buildup of PI, we created List Type as our within-subjects factor by collapsing list number across the two blocks. Note that we elected to block lists in an attempt to increase power and reliability. Consistent with prior research (Wickens, Born, & Allen, 1963; Wixted & Rohrer, 1993), a substantial effect of PI was observed. Figure 1 demonstrates the main effect of List Type,  $F(2, 318) = 482.97, MSE = .81, p < .001$ , partial  $\eta^2 = .75$ , indicating proportion of correctly recalled items decreased as PI accrued across lists.

Additionally, a main effect of WMC was observed,  $F(1, 159) = 3.90, MSE = 2.06, p = .05$ , partial  $\eta^2 = .02$ , suggesting low-WMC individuals recalled lower proportions of correctly recalled items in general when compared with people with high WMC ( $r = .16, p = .05$ ). Consistent with Kane and Engle (2000; see Experiment 2), the ANCOVA also yielded a significant interaction between List Type and WMC,  $F(2, 318) = 3.62, MSE = .81, p < .05$ , partial  $\eta^2 = .02$ , indicating that low-WMC individuals were more prone to the

**Table 1** Descriptive statistics for all measures

Measure	<i>N</i>	<i>M</i>	<i>SD</i>	Skew	Kurtosis	Reliability
Ospan	161	60.63	9.73	-1.31	2.10	.79
Symspan	161	29.19	7.09	-.49	-.33	.77
Rspan	161	57.43	11.83	-.94	.92	.78
Accuracy	161	.58	.08	.04	-.44	.61
Mean Latency (s)	161	17.00	2.69	-.01	.43	.56
Mean IRT (s)	161	3.87	.81	.79	1.31	.46
Total Repetitions	161	.91	1.37	4.17	28.26	.51
Total Intrusions	161	2.30	2.11	2.62	14.17	.47

detrimental effects associated with the buildup of PI. No significant differences between high WMC and low-WMC individuals appeared in proportion of correctly recalled items before the accumulation of PI or upon release from PI (i.e., List Type 1), but recall performance for low-WMC individuals was worse on remaining lists due to a quicker decline in recall performance as PI accumulated (see Fig. 1).

### Recall latency

Similar to the recall accuracy analysis, we submitted mean recall latency to a repeated-measures ANCOVA, with List Type as a within-subjects factor and WMC as a covariate. Results revealed a main effect of List Type,  $F(2, 278) = 17.99, MSE = 10.22, p < .001$ , partial  $\eta^2 = .12$ , suggesting mean recall latency increased as PI accrued across lists. There was also a main effect of WMC,  $F(1, 139) = 14.73, MSE = 20.84, p < .001$ , partial  $\eta^2 = .10$ , demonstrating low-WMC individuals recalled items more slowly overall than did high-WMC individuals ( $r = -.25, p = .001$ ). Furthermore, a significant List Type  $\times$  WMC interaction,  $F(2, 278) = 3.69, MSE = 10.22, p < .05$ , partial  $\eta^2 = .03$ , indicated that the differences between high-WMC and low-WMC individuals in mean recall latency became more extreme as PI increased across lists (see Fig. 2).

## Discussion

The results from Experiment 1 were relatively straightforward in demonstrating distinct PI effects, in which recall latency increased and recall accuracy decreased across lists. The slowing of retrieval is indicative of a less focused search set composed of excess irrelevant information (Wixted & Rohrer, 1993), whereas the decline in accuracy performance suggests an inability to recover target items. Critically, the accumulation of PI differentially impacted these aspects of the memory search for high-WMC and low-WMC individuals. That is, the buildup of PI impaired both accuracy performance (see Fig. 1) and recall latency (see Fig. 2) for low-WMC individuals to a greater extent than what was observed for high-WMC

<sup>2</sup> To investigate the possibility that the high kurtosis for repetitions and intrusions influenced our results, we first compared our Pearson product-moment correlation coefficients to a correlation coefficient less sensitive to nonnormally distributed data: the Spearman rank correlation coefficient (see de Winter, Gosling, & Potter, 2016). Importantly, while the strength of some of the correlations somewhat increased (e.g., Pearson's  $r$  between WMC and intrusions was  $-.14$ , but Spearman's  $r = -.17$ ), the overall pattern of results concerning correlations with these variables remained the same in both experiments. We also transformed the data in both experiments by employing a square root transformation. Despite correcting for skew and kurtosis (all skew values  $< .85$ ; all kurtosis values  $< .77$ ), rerunning the analyses with the transformed repetition and intrusion data also yielded similar results. Because the interpretability of our results did not change, we continued to use the raw data in our analyses.

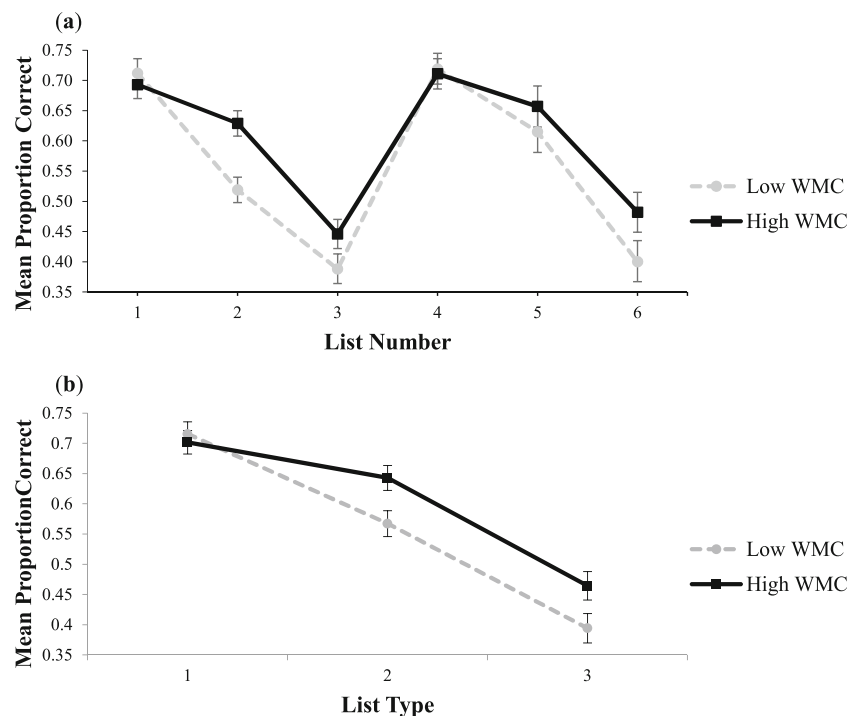
**Table 2** Correlations among all measures

Measure	1	2	3	4	5	6	7	8	9
1. WMC	—								
2. Ospan	.83***	—							
3. Symspan	.75***	.41***	—						
4. Rspan	.82***	.57***	.39***	—					
5. Accuracy	.16 <sup>+</sup>	.14	.04	.19*	—				
6. Mean Latency	-.25**	-.16*	-.21**	-.23**	-.09	—			
7. Mean IRT	-.17*	-.12	-.12	-.18*	-.25**	.53***	—		
8. Total Rep.	.14	.10	.09	.13	.13	.17*	-.09	—	
9. Total Intru.	-.14	-.07	-.07	-.20*	.00	.13	-.10	.05	—

Note. <sup>+</sup>  $p = .05$ ; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ ; Rep. = repetitions; Intru. = intrusions (total previous list intrusions + total extra list intrusions)

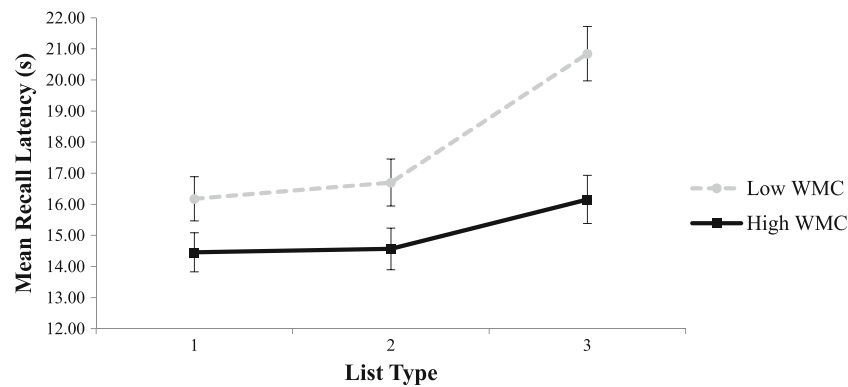
individuals. Thus, consistent with Kliegl et al. (2015), reducing temporal list discrimination via the buildup of PI successfully exaggerated WMC-related differences in search set size. PI seemingly had a larger effect on the size of the search set for low-WMC individuals because these individuals rely on noisier temporal-contextual cues at retrieval (Unsworth, 2007; Unsworth & Engle, 2007; Unsworth et al., 2012), and PI is one circumstance in which previously learned material is more easily incorporated into the search set.

It is also important to note that our recall accuracy findings are inconsistent with Kliegl et al. (2015) but do replicate Kane and Engle (2000; see Experiment 2). Namely, no WMC-related differences in recall accuracy were observed when PI was minimal (i.e., on List 1), but differences between high-WMC and low-WMC individuals emerged and became significantly larger as PI increased. Considering the fact that we adapted our PI buildup task from Kane and Engle (2000) and that both studies recruited large sample sizes ( $N$  in the



**Fig. 1** **a** Mean proportion of correctly recalled items as a function of list number and WMC. Note that list 1 represents no PI, whereas List 4 represents a release from PI (i.e., semantic category switch). **b** Mean proportion of correctly recalled items as a function of WMC span group

and List Type—1 being the first time encountering semantic category (when PI is lowest) and 3 being the third time encountering semantic category (when PI is highest). Error bars reflect one standard error of the mean



**Fig. 2** Mean recall latency as a function of WMC span group and List Type. Error bars reflect one standard error of the mean

present experiment = 161;  $N$  in Kane and Engle's, 2000, Experiment 2 = 192), it is unsurprising that these experiments produced similar WMC-related PI effects in recall accuracy. Therefore, the discrepancy in our results from Kliegl et al.'s (2015) results may be attributed to their use of a slightly different procedure, or, more plausibly, their smaller sample size of 60 participants.

## Experiment 2

The purpose of Experiment 2 was to better elucidate whether it is possible to eliminate WMC-related differences in memory search abilities by modifying the relevance of previously learned material. To assess this question, we provided participants with additional list learning experience via multitrial free recall. Multitrial free recall involves learning the same list of words over multiple study test trials, with the order of word presentation within a list varying from trial to trial. Importantly, research (Dunlosky & Salthouse, 1996; Sternberg & Tulving, 1977; Tulving, 1966) utilizing multitrial free recall paradigms has repeatedly revealed that the number of words recalled increases across study test trials, meaning participants learn the list as a function of practice. This increase in recall performance across trials is believed to be partially due to what researchers refer to as subjective organization, the tendency of recall output to become progressively more fixed across trials (i.e., certain words are grouped and recalled together on successive trials; Tulving, 1962).

It remains to be seen, though, how multitrial free recall influences other aspects of the search (recall latency and IRTs), and whether these effects differ as a function of WMC. If WMC-related differences in search set size are due to differences in the effectiveness of one's retrieval cues (Unsworth, 2007; Unsworth & Engle, 2007; Unsworth et al., 2012), then increased practice learning the same word list may result in low-WMC individuals searching LTM as efficiently as high-WMC individuals in some, but not all, aspects of the memory search. To better understand this hypothesis, it is first

vital to note that in traditional free recall experiments, each list contains different to-be-remembered items. As was the case in Experiment 1, items from previously learned lists become irrelevant over the course of learning because individuals are instructed to only recall items from the current list. However, in cases of multitrial free recall, previously learned items remain relevant because the same items are being presented across trials. Thus, it is likely that a new copy of each word is created every time the word list is presented (see Bower, 1967; Hintsman, 1988; Unsworth, 2015) and/or when the words are recalled, resulting in an enlarged search set. Note that within search models there are potentially two ways that copies can influence the search. One possibility is that the inclusion of copies may influence the overall size of the search set, but this has no effect on recall latency (e.g., Wixted, Ghadisha, & Vera, 1997). In this case, it is assumed that once an item is sampled and recalled, that item and any of its copies are marked as being recalled and thus will not be recalled again. For example, consider a list of 10 items presented once. The probability of sampling any one item will be  $1/10$ . Similarly, if a list of 10 items are presented twice the probability of selecting a given item should be  $2/20$  or  $1/10$ . Thus, in this case recall latency should be the same for both conditions. Critically, this situation assumes that participants are treating copies as the same and recognize them as being the same. Another possibility is that participants do not fully recognize the copy as a repeated item, but rather treat both items as being distinct. This would result in an overall increase in the size of the search set and increase in recall latency because both the original item and the copy would be considered as unique with a probability of sampling any one item being  $1/20$  (similar to what is seen with list-length effects; Rohrer & Wixted, 1994). Furthermore, this should result in an increase in repetition errors as participants are likely to recall both items (the original and the copy). Support for the latter notion arises from encoding manipulations in which to-be-remembered items are presented in a spaced format (see Unsworth, 2015). Spaced learning conditions are like multitrial free recall in the sense that the same word list is presented once and then presented

again (the difference being that participants are tested for recall multiple times in multitrial free recall). Unsworth (2015) revealed that under spaced learning conditions, both recall accuracy and recall latency increased as a function of list presentation, whereas IRTs decreased (i.e., became faster). Importantly, results were similarly interpreted with respect to multiple copies models of memory (see Bower, 1967; Hintsman, 1988), such that each word presentation within the spaced learning condition was thought to result in the creation of a new distinct copy of the target item (see also Ozier, 1980).

If low-WMC individuals are more likely to incorporate previously learned material into their search sets, conditions of multitrial free recall may result in the inclusion of multiple copies of a single target item (opposed to previous list intrusions in normal free recall conditions with and without a PI manipulation). While the accumulation of more target-item copies within the search set should still result in slower recall latencies for low-WMC individuals, other aspects of the search may actually be facilitated. That is, with multiple copies of a single target item being included within a search set, one may encounter an easier time locating said target item within the search set as the search is underway (i.e., faster IRTs), as well as a greater likelihood of sampling said target item (resulting in improved recall accuracy).

Taken altogether, if low-WMC individuals struggle to confine the search to current-list items, multitrial free recall learning should result in the following pattern of results for these individuals: increased recall accuracy and recall latency, but decreased IRTs. In contrast, since people with high WMC are better able to focus the search to current-list items via their use of more effective retrieval cues, we would not expect these individuals to include as many target-item copies within their search sets. Thus, while high-WMC individuals may still benefit from repeated study test trials, aspects of their memory search should not be as affected as is the case for low-WMC individuals. In other words, with increased list-learning experience, low-WMC individuals should be able to search LTM as efficiently as their high-WMC counterparts.

## Method

### Participants and procedure

Participants were 127 undergraduate students (59% female) recruited from the human subject pool at the University of Oregon, and all participants were between the ages of 18 and 30 years ( $M = 19.76$  years,  $SD = 2.01$ ). Again, nonnative English speakers were excluded from the study, and all eligible participants were compensated with course research credit for participation. The materials and procedures for the WMC tasks were identical to those reported in Experiment 1, and each participant was similarly tested individually in a single

laboratory session lasting approximately 2 hours. As part of the 2-hour experimental session, participants first completed the WMC tasks, followed by a multitrial free recall task. Participants were also asked to complete other tasks not of interest to the present investigation. Again, these tasks are not reported.

### WMC tasks

See Experiment 1.

### Multitrial free recall task

The multitrial free recall task was analogous to the delayed free recall task described in Experiment 1, with the exception that participants received eight trials of the same word list composed of 10 unrelated nouns. Moreover, the order in which these words appeared across presentations was randomized. Words were presented on computer for 1 second each, and following presentation of the last word within each list, a 16-second distractor task began that required participants to verbally report a series of eight three-digit numbers in descending order (adapted from Rohrer & Wixted, 1994). Each three-digit string was presented onscreen for 2 seconds. After the distractor task, three question marks appeared in the center of the screen to prompt participants to recall as many words as possible within a 45-second window. Participants typed their responses in any order they wished and pressed “enter” after each word, thereby clearing the screen. RTs were again measured with respect to when participants pressed enter after the word was typed. The dependent variables of interest were total proportion of correctly recalled items for each list, as well as recall latency and IRTs. Total repetitions and total extra-list intrusions were likewise recorded.

We also measured subjective organization for each participant by calculating bidirectional paired frequency (PF) scores for every consecutive pair of trials.<sup>3</sup> Note that Pair 1 represents the PF score for Trial 1 and Trial 2, Pair 2 represents the PF

<sup>3</sup> The equation for calculating bidirectional paired frequency scores (see Sternberg & Tulving, 1977) is the following:  $O(ITR2) - E(ITR2) = O(ITR2) - (2c(c-1))/hk$ .  $O(ITR2)$  is the number of pairs of words recalled on Trial  $t$  and Trail  $t + 1$  in adjacent output positions in either forward or backward order (hence, the name bidirectional), whereas  $E(ITR2)$  reflects the expected number of paired words. The letter  $c$  is the number of common words recalled on both trials,  $h$  is the number of words recalled on Trial  $t$ , and  $k$  is the number of words recalled on Trial  $t + 1$ . As a demonstration: Imagine the order of word presentation for List 1 is *hook, fan, rake, rib, yard, sock, zoo, ditch, sink, and vest*, and the order of word presentation for List 2 is *ditch, vest, rake, sink, fan, zoo, sock, rib, hook, and yard*. The participant recalls the following six words on List 1: *hook, fan, rake, yard, zoo, and sink*. On List 2, the same participant recalls the following eight words: *hook, fan, rake, ditch, yard, zoo, vest, and sock*. With respect to  $O(ITR2)$ , notice that the following pairs of words appeared in adjacent positions across both lists: *hook-fan, fan-rake, and yard-zoo*, yielding a score of 3.  $E(ITR2)$  would be calculated as  $(2(5)(5-1))/(6(8))$ . Taking the whole equation into consideration, this participant’s PF score for Trials 1 and 2 would be 2.17.



score for Trial 2 and Trial 3, and so forth. We selected bidirectional PF as our assessment of subjective organization because, according to Sternberg and Tulving (1977), it is the most reliable and empirically valid measure that correlates highly with free recall and other indices of subjective organization. Indeed, more recent research interested in subjective organization (e.g., Davis et al., 2003; Harrison, 2014; Weldon & Bellinger, 1997) continues to use bidirectional PF as the measure of choice.

## Results

Similar to Experiment 1, a WMC composite score was computed for each participant via mean standardization of subjects' complex span scores, and WMC was always treated as a continuous variable. For all of the following analyses, we submitted various dependent variables to a series of repeated-measures ANCOVAs. In each analysis, list is the within-subjects factor and WMC is the covariate (see Table 3 for descriptive statistics and Table 4 for correlations among all measures).

### Proportion correct recall

For proportion of correctly recalled items, results revealed a main effect of trial,  $F(7, 875) = 170.83$ ,  $MSE = .01$ ,  $p < .001$ , partial  $\eta^2 = .58$ , demonstrating a steep increase in recall performance across the first few trials, followed by a more gradual growth in performance across the remaining trials. One finding in stark contrast to Experiments 1 is that the between-subjects effect of WMC was nonsignificant,  $F = 1.20$ ,  $p > .27$ , meaning high-WMC and low-WMC individuals did not significantly differ in recall accuracy in general. Critically, though, the interaction between trial and WMC was significant,  $F(7, 875) = 2.13$ ,  $MSE = .01$ ,  $p < .05$ , partial  $\eta^2 = .02$ . Figure 3 reveals that differences between high-WMC and low-WMC individuals in recall accuracy did exist on early study test trials, such that higher WMC was related to a greater proportion of correctly recalled items on Trial 1 ( $r = .18$ ,  $p < .05$ ). However, differences between high-WMC and low-WMC individuals continued to diminish across trials, with no observable differences in proportion of correctly recalled items by Trial 4 ( $r = .06$ ,  $p > .52$ ).

### Interresponse times

In regard to mean IRTs, analyses yielded a between-subjects effect of WMC,  $F(1, 123) = 3.96$ ,  $MSE = 4.46$ ,  $p < .05$ , partial  $\eta^2 = .03$ , suggesting higher WMC was related to faster IRTs overall ( $r = -.18$ ,  $p < .05$ ).<sup>4</sup> There was also a main effect of

<sup>4</sup> Two participants were excluded from this analysis because they did not have IRT data for the first two trials.

**Table 3** Descriptive statistics for all measures

Measure	<i>N</i>	<i>M</i>	<i>SD</i>	Skew	Kurtosis	Reliability
Ospan	127	57.59	12.03	-1.12	.84	.76
Symspan	127	29.67	6.21	-.47	-.08	.58
Rspan	127	52.41	12.98	-.68	.47	.79
Accuracy	127	.76	.12	-.68	.31	.88
Mean Latency (s)	127	13.74	3.29	.44	-.10	.86
Mean IRT (s)	127	3.07	.72	.69	.77	.75
Total Repetitions	127	9.33	8.25	2.01	5.70	.84
Total ELI	127	3.58	5.45	2.51	7.16	.84
Mean SO	127	.54	.60	1.29	3.55	.53

*Note.* ELI = extra list intrusions; SO = subjective organization (bi-directional paired frequency)

trial,  $F(7, 861) = 21.02$ ,  $MSE = 1.10$ ,  $p < .001$ , partial  $\eta^2 = .15$ , demonstrating a decline in IRTs with each subsequent study—test trial. Of particular importance, a significant interaction between trial and WMC,  $F(7, 861) = 4.64$ ,  $MSE = 1.10$ ,  $p < .001$ , partial  $\eta^2 = .04$ , further revealed that WMC-related differences in IRTs were reduced and eventually eliminated with each study test trial. More specifically (see Fig. 4), high WMC and low-WMC individuals tended to display faster IRTs across the first 5 trials, but this high level of performance remained at a stable level across the remaining trials. Moreover, while no WMC-related differences were detected on Trial 1 ( $r = .03$ ,  $p > .72$ ), WMC-related differences in IRTs did emerge on Trials 2 ( $r = -.36$ ,  $p < .001$ ) through 4 ( $r = -.18$ ,  $p < .05$ ). Following Trial 4, though, these differences soon diminished and became nonexistent with each subsequent trial (all  $ps > .53$ ). Thus, increased practice with the same word list served to equate the memory search abilities of high-WMC and low-WMC individuals, particularly in terms of IRTs.<sup>5</sup>

### Recall latency

With respect to mean recall latency, a main effect of trial,  $F(7, 861) = 3.58$ ,  $MSE = 12.10$ ,  $p = .001$ , partial  $\eta^2 = .03$ , indicated that recall latencies tended to increase (i.e., become slower) with each trial. The ANCOVA also revealed a between-subjects effect of WMC,  $F(1, 123) = 9.88$ ,  $MSE = 79.30$ ,  $p < .01$ , partial  $\eta^2 = .07$ , suggesting higher WMC was associated with faster recall latencies overall ( $r = -.26$ ,  $p < .01$ ). Moreover, a significant interaction between trial and WMC,  $F(7, 861) = 2.99$ ,  $MSE = 12.10$ ,  $p < .01$ , partial  $\eta^2 = .02$ , showed a minimization in recall latency differences between high WMC and low-WMC individuals as trials increased (see Fig. 5). That is, despite no observable WMC-related

<sup>5</sup> Another potential reason for the change in IRTs across lists could be that as recall is becoming more organized the search is transitioning from sampling-with-replacement to sampling-without-replacement, which would lead to faster IRTs overall. Future research is needed to examine this possibility.

Table 4 Correlations among all measures

Measure	1	2	3	4	5	6	7	8	9	10
1. WMC	—									
2. Ospan	.86***	—								
3. Symspan	.79***	.52***	—							
4. Rspan	.82***	.60***	.44***	—						
5. Accuracy	.10	.04	.12	.09	—					
6. Mean latency	-.26**	-.19*	-.14	-.31***	.15	—				
7. Mean IRT	-.16	-.07	-.08	-.24**	-.25**	.57***	—			
8. Total rep.	-.20*	-.17 <sup>+</sup>	-.18*	-.14	.19*	.34***	-.03	—		
9. Total ELI	.05	.04	-.01	-.08	-.31***	.15	.12	.23**	—	
10. Mean SO	.11	.06	.16	.06	.32***	-.09	-.24**	-.21*	-.14	—

Note. <sup>+</sup>  $p = .06$ ; \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ ; ELI = extra list intrusions; SO = subjective organization (bi-directional paired frequency)

differences in recall latency on Trial 1 ( $r = .01$ ,  $p > .93$ ), differences were observed across the majority of remaining trials. The largest WMC differences in recall latency emerged on Trials 2 ( $r = -.35$ ,  $p < .001$ ) and 3 ( $r = -.29$ ,  $p = .001$ ). Following Trial 4 ( $r = -.21$ ,  $p < .05$ ), these differences in recall latency became less extreme and in some cases even disappeared ( $p$ s for Trials 5 through 7  $> .09$ ). However, WMC-related differences soon reemerged on the last trial ( $r = -.22$ ,  $p < .05$ ). Taken altogether, these results suggest that even when the same word list is studied and recalled across numerous trials, both low-WMC and high-WMC individuals continue to accumulate larger search sets with each list repetition. Critically, it appears that this increased practice served to diminish, but not eliminate, recall latency differences between high-WMC and low-WMC individuals.

### Repetitions

It seems conceivable that if individuals include multiple copies of target items within the search set, then this may, in turn, be reflected in recalling more repetitions across trials. In support of this notion, the repeated-measures ANCOVA analyzing repetitions showed a main effect of trial,  $F(7, 875) = 12.10$ ,  $MSE = 1.41$ ,  $p$

$< .001$ , partial  $\eta^2 = .09$ , suggesting repetitions did increase as trials increased. An examination of Fig. 6 reveals a more pronounced increase in repetitions across the first four trials followed by a plateau in repetitions across the remaining trials. Results also suggested a between-subjects effect of WMC,  $F(1, 125) = 5.02$ ,  $MSE = 8.24$ ,  $p < .05$ , partial  $\eta^2 = .04$ , indicating less repetition errors overall were related to higher WMC ( $r = -.20$ ,  $p < .05$ ). Of note, the interaction between trial and WMC was nonsignificant,  $F < 1$ ,  $p > .60$ , meaning (1) low-WMC individuals' tendency to emit more repetitions than high-WMC individuals did not change as a function of trial, and (2) there was an increase in repetitions across trials for both high-WMC and low-WMC individuals. While it is possible that people with low WMC recalled more repetitions because these individuals are more likely to forget what they have recently recalled (for review on WMC-related differences in monitoring abilities, see Unsworth & Brewer, 2010), we believe these results are consistent with the notion that the increase in repetitions, or copies of target items, across trials likely contributed to the increase in search set size across trials ( $r$  between mean recall latency and mean repetitions = .34,  $p < .001$ ).

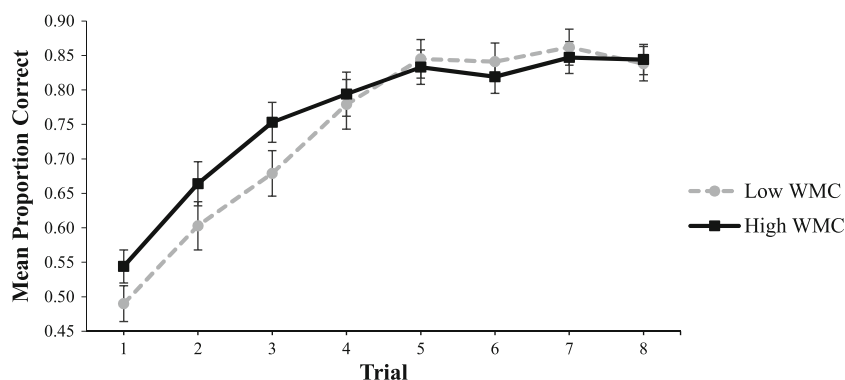


Fig. 3 Mean proportion of correctly recalled items by trial and WMC. Error bars reflect one standard error of the mean

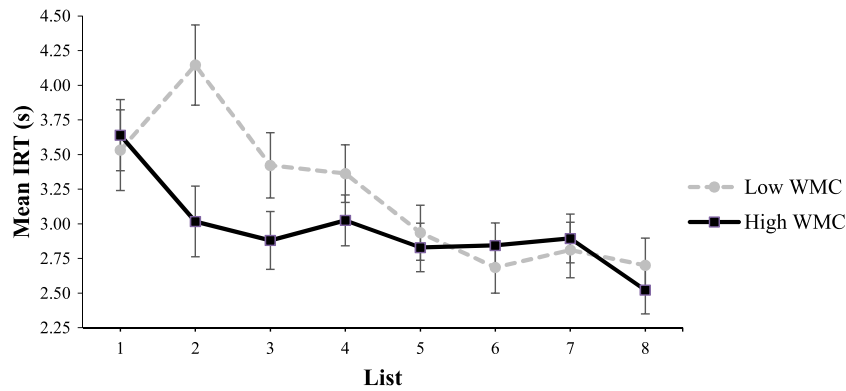


Fig. 4 Mean IRT as a function of trial and WMC. Error bars reflect one standard error of the mean

**Subjective organization**

Since recent research (Harrison, 2014) suggests fundamental differences in WMC are positively related to mean subjective organization when collapsing across all study test trials, we sought to likewise investigate this relation. Subjective organization was measured by calculating the bidirectional PF score for each successive pair of trials for each participant. The repeated-measures ANCOVA, with subjective organization as the dependent variable, revealed a main effect of pair,  $F(6, 738) = 6.22, MSE = 1.80, p < .001, \text{partial } \eta^2 = .05$ , indicating an increase in subjective organization across each successive pair of trials (see Fig. 7). In addition, the between-subjects effect of WMC was nonsignificant,  $F = 1.93, p > .16$ , suggesting high-WMC and low-WMC individuals did not significantly differ in the extent to which recall output became increasingly stereotyped. The interaction between pair and WMC was also nonsignificant,  $F < 1, p > .80$ .

**Discussion**

The results from Experiment 2 revealed that when previously learned material remains relevant through the course of learning, it is possible to equate some, but not all, aspects of the memory search for low-WMC and high-WMC individuals. Specifically, while high-WMC individuals displayed better recall accuracy than did low-WMC individuals across the first

few study test trials, these differences became nonexistent by Trial 4 (see Fig. 3). Moreover, while no WMC-related differences in IRTs were observed on the first study-test trial, substantial IRT differences emerged on the second study-test trial. Similar to recall accuracy, the IRT differences between high-WMC and low-WMC individuals eventually disappeared following the fourth study-test trial (see Fig. 4). Together, these results suggest practice served to help low-WMC individuals more than high-WMC individuals by facilitating low-WMC individuals’ ability to recover additional target items more quickly as the search was underway. However, despite no WMC-related differences in recall latency on study-test Trial 1, recall latency increased for all individuals across the remaining study test trials, and this effect was larger for people with low WMC (see Fig. 5). This suggests that over the course of repeated study test trials, low-WMC individuals still accumulated larger search sets than their high-WMC counterparts.

The overall pattern of results (improved recall accuracy and faster IRTs, but still slower recall latencies for low-WMC individuals) supports the view that low-WMC individuals engage in less efficient memory searches under normal learning conditions because of their inability to use precise temporal-contextual cues at retrieval. Before prior-list-learning experience could influence the size of one’s search set (i.e., List 1), no WMC-related differences were observed in recall latency or IRTs. Across the remaining study test trials, though, list-learning experience served to produce multiple copies of

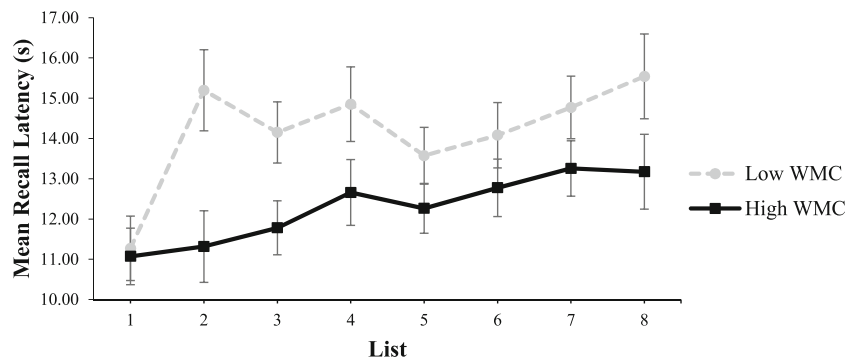
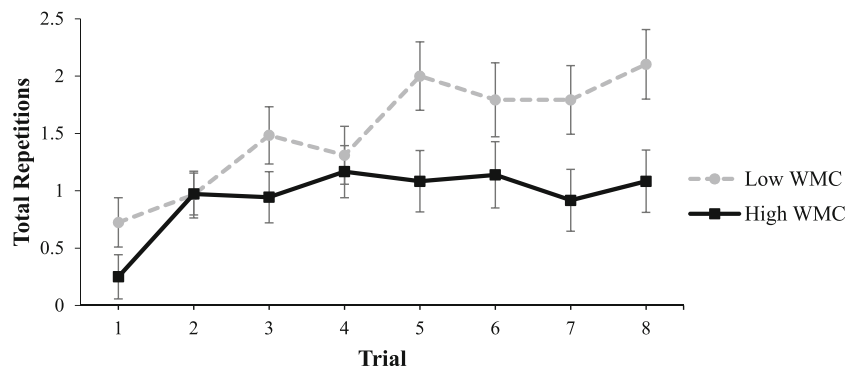


Fig. 5 Mean recall latency as a function of trial and WMC. Error bars reflect one standard error of the mean



**Fig. 6** Total repetitions as a function of trial and WMC. Error bars reflect one standard error of the mean

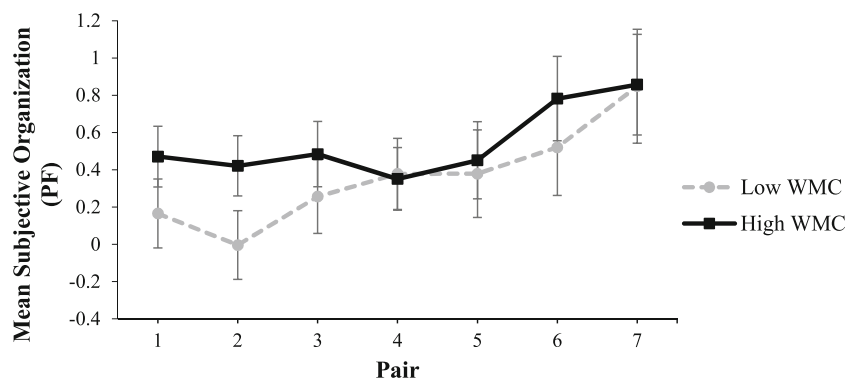
target items within the search set, instead of previous list intrusions. This occurred because in multitrial free recall, words learned on previous trials remain relevant throughout the course of learning, meaning each list presentation results in the creation a new target item copy. If individuals, such as those with low WMC, use vague cues that interfere with their ability to discriminate previous-list items from current-list items, then conditions of multitrial free recall seemingly result in culminating larger search sets composed of multiple target-item copies. Assuming that conditions of multitrial free recall serve to produce multiple distinct copies of target items, repetitions should theoretically increase as a function of study test trial. Indeed, the results from Experiment 2 revealed that repetitions increased across trials and were negatively correlated with WMC ( $r = -.20, p < .05$ ), yet positively correlated with recall accuracy ( $r = .19, p < .05$ ) and recall latency ( $r = .34, p < .001$ ).

In sum, we were not able to eliminate WMC-related differences in search set size. Across trials, low-WMC individuals were more likely to include items from previous trials in their search sets, resulting in overall larger search sets made up of more target-item copies (akin to repetitions). Fortunately, changing the relevance of previously learned material actually benefited other aspects of the memory search for low-WMC individuals, such that practice with the same word list served to equate recall accuracy and IRTs for low-WMC and high-WMC individuals. Furthermore, it appears that subjective

organization can be ruled out as a factor responsible for WMC-related differences in search-efficiency processes. That is, while subjective organization was positively associated with recall performance ( $r = .32, p < .001$ ) and increased across study test trials (see Fig. 6), it was not related to differences in WMC ( $r = .11, p = .21$ ).

### General discussion

In two experiments, using variants of a delayed free recall task, we examined individual differences in WMC and the ability to efficiently search the contents of LTM. Given the notion that higher WMC is associated with faster recall latencies (Unsworth, 2007, 2009a, 2009b; Unsworth & Engle, 2007; Unsworth et al., 2012), Experiment 1 sought to test whether the buildup of PI could exaggerate WMC-related differences in search set size. Results revealed that while the buildup of PI resulted in longer recall latencies and decreased recall accuracy for all participants (see Wickens et al., 1963; Wixted & Rohrer, 1993), the PI manipulation was successful in interfering with both these features of the memory search more so for low-WMC individuals than for high-WMC individuals. Under conditions of PI, it becomes progressively more difficult to discriminate current-list items from previous-list items. Presumably, low-WMC individuals were especially sensitive to these effects because they utilize noisier



**Fig. 7** Subjective organization scores by pair and WMC. Error bars reflect one standard error of the mean



temporal-contextual cues at retrieval. By the use of these cues, low-WMC individuals are less able to filter out irrelevant items and focus the search to current-list items, resulting in worse recall accuracy and slower recall latencies (Unsworth, 2009a, b, 2015, 2016). If low-WMC individuals search through more items at retrieval because they are more likely to incorporate irrelevant information—such as previously learned material—into their search sets, one may then wonder how other aspects of the search (e.g., recall accuracy and IRTs) may be influenced if previously learned information remains relevant throughout the course of learning. In an attempt to answer this question, Experiment 2 employed a multitrial delayed free recall task. Results demonstrated that with increased list-learning practice with the same word list, low-WMC individuals were able to search LTM as efficiently as their high-WMC counterparts. Namely, despite the finding that low-WMC individuals continued to search through more items at retrieval across all trials (relative to high-WMC individuals), high-WMC individuals no longer outperformed their low-WMC counterparts in recall accuracy or IRTs by the fourth study test trial.

Collectively, the results from Experiment 1 and Experiment 2 suggest that low-WMC individuals undergo a less focused search through LTM. That is, people with low WMC search through larger search sets because these individuals use noisier temporal-contextual cues to guide retrieval. Under normal free recall conditions in which participants encounter only one study test trial per list, use of these noisy cues results in accumulating larger search sets composed of more irrelevant items, items that typically share more contextual features with current list items. But it appears that the reliance on noisy retrieval cues can actually prove to be advantageous in other circumstances. Specifically, under conditions providing multiple study test trials with the same to-be-remembered material, a new distinct copy of each word is likely created in LTM every time the word list is presented (see Bower, 1967; Hintsman, 1988; Unsworth, 2015). Therefore, by use of vague temporal-contextual cues, low-WMC individuals likely include more target-item copies within their search sets under these circumstances. While the inclusion of multiple target-item copies still results in accumulating larger search sets for these individuals, the inclusion of more target-item copies seemingly leads to increased performance in other aspects of the search. That is, there is an increased likelihood that a correct item is sampled and recovered, as well as an improved ability to swiftly locate target items within the search set as the search is underway. Since people with high WMC utilize more precise temporal-contextual cues, these individuals likely do not include as many target-item copies within their search sets, meaning aspects of their memory search are not as affected by practice as is the case for low-WMC individuals.

Nonetheless, it is important to note that rather than WMC-related differences in recall latency reflecting variability in the

breadth of search (attributed to one's use of retrieval cues), an alternative account suggests that WMC-related differences in search-efficiency abilities may be largely driven by differences in speed of processing. The speed-of-processing view is most prominent in the developmental literature (Hale, 1990; Salthouse, 1996), whereby speed of processing is believed to be the primary source of variation responsible for age-related differences in cognitive ability. In regard to variation in general cognitive ability among adults, these theories would posit that individuals with greater cognitive ability have quicker IRTs and recall latencies because they are able to process information at a faster rate (Jensen, 1998). Indeed, an examination of Fig. 2 in Experiment 1 reveals that high-WMC and low-WMC individuals differed in mean recall latency on List Type 1 (i.e., at List 1 when minimal PI was present and at List 4 when PI was released). If virtually no PI was present on these lists, yet low-WMC individuals were still slower to recall items, it seems plausible that differences in speed of processing could explain these results. Nonetheless, we believe Experiment 2 rules out this possibility. Namely, if speed of processing were the underlying factor responsible for the slowed retrieval characteristic of people with low WMC, a different pattern of results should have arisen under conditions of multitrial free recall: Even with practice, high-WMC and low-WMC individuals should have always differed in time-sensitive aspects of the memory search. Because practice actually served to equate IRTs for low-WMC and high-WMC individuals, it appears that speed of processing can be ruled out as a factor responsible for WMC-related differences in search-efficiency processes.

### Limitations

A possible limitation with the current study concerns the list length used in our multitrial free recall procedure. We provided participants with a word list containing 10 nouns, whereas other studies employing multitrial free recall have typically used lists containing between 15 (Dunlosky & Salthouse, 1996) and 35 (Harrison, 2014) nouns. Therefore, it seems within reason to speculate that participants in Experiment 2 of the present study may have been performing at ceiling with respect to recall accuracy (see Fig. 3). Indeed, recall accuracy appeared to level off around 85%. While future work should aim to replicate and extend our results with longer word lists, we do not believe the presence of a ceiling effect in recall accuracy influences or outweighs the other important aspects of our analyses.<sup>6</sup> The finding that participants were unable to

<sup>6</sup> To investigate whether the 45-second time limit at recall contributed to our 85% asymptotic recall, we plotted the cumulative recall curves for each trial. Importantly, while the time taken to recall items increased with each list repetition, participants reached asymptotic recall around 35 seconds on List 8. This result suggests that the recall period was not too short for all items to be recalled.

recall all items by the last trial is consistent with prior work examining multitrial free recall performance. For instance, Dunlosky and Salthouse (1996) presented participants with a 15-item word list across five trials, revealing minimal improvement in recall accuracy between Trial 4 and Trial 5 (asymptote between 85% and 90%). Using a word list containing 19 total items, Tulving (1966) demonstrated that across 12 trials, participants plateaued between 80% (Trial 8) and 90% (Trial 12) correct. Harrison (2014) administered slightly longer word lists of 25 items across nine trials and similarly found an asymptote occur around 75% correct. Moreover, a primary question of interest in Experiment 2 was whether we could equate the speed with which low-WMC and high-WMC individuals search for items in LTM. Figure 5 reveals that the fastest IRTs for both high-WMC and low-WMC individuals were around 2,500 ms. Unsworth and Engle (2007) demonstrated that it is possible to observe IRTs as fast as 1,000 ms for all individuals. Thus, while participants in Experiment 2 of the present study may have performed at ceiling in recall accuracy, these individuals did not reach a floor effect in IRTs.

## Conclusion

All things considered, the overall pattern of results is largely consistent with research suggesting that when compared with high-WMC individuals, low-WMC individuals are less able to engage in an efficient memory search because of their reliance on noisy temporal-contextual cues at retrieval (Unsworth, 2007; Unsworth & Engle, 2007; Unsworth et al., 2012). The reliance on imprecise retrieval cues results in overall larger search sets composed of more irrelevant information, resulting in a slower, less accurate memory search under normal free recall conditions. Critically, though, the use of noisy retrieval cues actually proves to be advantageous in circumstances whereby previously learned material remains relevant over the course of learning, to the point where low-WMC individuals are able to search LTM just as efficiently as high-WMC individuals.

## References

- Bower, G. (1967). A multicomponent theory of the memory trace. *Psychology of Learning and Motivation* (Vol. 1). New York, NY: Academic Press.
- Craik, F.I.M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409–422). Amsterdam, The Netherlands: Elsevier.
- Davelaar, E. J., & Raaijmakers, J. G. W. (2012). Human memory search. In P. M. Todd, T. Hills & T. Robbins (Eds.), *Cognitive search: Evolution, algorithms, and the brain*. (pp. 177–193). Cambridge, MA: MIT Press.
- Davis, H. P., Small, S. A., Stern, Y., Mayeux, R., Feldstein, S. N., & Keller, F. R. (2003). Acquisition, recall, and forgetting of verbal information in long-term memory by young, middle-aged, and elderly individuals. *Cortex*, *39*, 1063–1091.
- de Winter, J.C.F., Gosling, S.D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological Methods*, *21*, 273–290.
- Dunlosky, J., & Salthouse, T. A. (1996). A decomposition of age-related differences in multitrial free recall. *Aging, Neuropsychology, and Cognition*, *3*, 2–14.
- Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development*, *61*, 653–663.
- Harrison, T. L. (2014). The role of working memory burden and fluid intelligence in the organization of memory (Unpublished doctoral thesis). Georgia Institute of Technology, Atlanta, Georgia.
- Hintsman, D. L. (1988). Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *4*, 528–551.
- Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.
- Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 336–358.
- Kliegl, O., Pastotter, B., & Bauml, K. H. T. (2015). The contribution of encoding and retrieval processes to proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1778–1789.
- Mickes, L., Seale-Carlisle, T. M., & Wixted, J. T. (2013). Rethinking familiarity: Remember/know judgments in free recall. *Journal of Memory and Language*, *68*, 333–349.
- Murdock, B. B., & Okada, R. (1970). Interresponse times in single-trial free recall. *Journal of Experimental Psychology*, *86*, 263–267.
- Ozier, M. (1980). Individual differences in free recall: When some people remember better than others. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 309–364). New York, NY: Academic Press.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 207–262). New York, NY: Academic Press.
- Rohrer, D. (1996). On the relative and absolute strength of a memory trace. *Memory & Cognition*, *24*, 188–201.
- Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition*, *22*, 511–524.
- Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*, 211–227.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.
- Salthouse, T. A. (2001). The broader context of Craik's self-initiated processing hypothesis. In M. Naveh-Benjamin, M. Moscovitch, & H. L. Roediger (Eds.), *Perspectives on human memory and cognitive aging: Essays in honour of Fergus Craik* (pp. 277–285). New York, NY: Psychology Press.
- Shiffrin, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of human memory* (pp. 375–447). New York, NY: Academic Press.
- Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review*, *76*, 179–193.
- Sternberg, R. J., & Tulving, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin*, *84*, 539–556.
- Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, *69*, 344–354.
- Tulving, E. (1966). Subjective organization and effects of repetition in multitrial free recall learning. *Journal of Verbal Learning and Verbal Behavior*, *5*, 193–197.
- Tulving, E. (1968). Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 2–36). Englewood Cliffs, NJ: Prentice Hall.

- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of psychology*, *53*, 1–25.
- Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, *64*, 49–60.
- Unsworth, N. (2007). Individual differences in working memory capacity and episodic memory: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 1020–1034.
- Unsworth, N. (2009a). Examining variation in working memory capacity and retrieval in cued recall. *Memory*, *17*, 386–396.
- Unsworth, N. (2009b). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition*, *37*, 837–849.
- Unsworth, N. (2010a). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence*, *38*, 255–267.
- Unsworth, N. (2010b). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. *Acta Psychologica*, *134*, 16–28.
- Unsworth, N. (2015). The influence of encoding manipulations on the dynamics of free recall. *Memory & Cognition*, *43*, 60–69.
- Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 50–61.
- Unsworth, N., & Brewer, G. A. (2010). Variation in working memory capacity and intrusions: Differences in generation or editing? *European Journal of Cognitive Psychology*, *22*, 990–1000.
- 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*, 104–132.
- Unsworth, N., & Engle, R.W. (2008). Speed and accuracy of accessing information in working memory: An individual differences investigation of focus switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 616–630.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498–505.
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*, 635–654.
- Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *Quarterly Journal of Experimental Psychology*, *65*, 2397–2410.
- Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1160–1175.
- Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *2*, 440–445.
- Wixted, J. T., Ghadisha, H., & Vera, R. (1997). Recall latency following pure- and mixed-strength lists: A direct test of the relative strength model of free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *23*, 523–538.
- Wixted, J. T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1024–1039.
- Wixted, J. T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review*, *1*, 89–106.