Manuscript accepted at *Journal of Experimental Psychology: Learning, Memory, and Cognition*

Where Working Memory Meets Long-Term Memory: The Interplay of List Length and Distractors on Memory Performance

Vanessa M. Loaiza

Sindre C. Halse

Word count: 14,634

Author Note

Vanessa M. Loaiza, Department of Psychology, University of Essex. Sindre C. Halse, Department of Psychology, University of Essex.

This work was supported, in part, by the University of Essex Department of Psychology's Research Promotion Fund awarded to Vanessa Loaiza.

We thank Geoff Ward for helpful comments on this work and Nathanael Knight for his assistance with data collection.

Correspondence concerning this article should be addressed to Vanessa M. Loaiza, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ. Email: v.loaiza@essex.ac.uk

The data and analysis scripts for the experiments are available at the Open Science Framework at: https://osf.io/b9evq/

Abstract

Word count: 237

Previous work regarding a counterintuitive benefit of increasing distractors on episodic long-term memory (LTM) has suggested that retrieval of memoranda in working memory (WM) after attention has been distracted may confer benefits to episodic LTM. The current study investigated two conceptions of how this may occur: either as an attentional refreshing of active memoranda within the focus of attention or as retrieval of a cohesive chunk of memoranda from outside the central component of WM. Given the literature suggesting that increasing the number of items to maintain in WM, or list length, incurs an attentional cost, the current study investigated whether increasing list length may reduce the beneficial impact of distractors on episodic LTM. In a series of three experiments we manipulated list length and the number of distractors following the memoranda in a Brown-Peterson-like-span task. Despite profound negative effects of list length and distractors on initial recall, the results indicated that list length did not interact with the beneficial effect of distractors on final free recall of the items. Furthermore, final free recall was consistent across serial position, in line with the view that all of the memoranda are retrieved as a chunk after each distractor. These findings emphasize the notion that recovering inactive information from outside of the central component of WM may impact its long-term retention. The theoretical implications regarding how retrieval may be a means by which LTM processes influence WM are discussed.

Keywords: working memory, episodic memory, long-term memory, list length effects, set size effects

How do humans maintain relevant information from moment to moment, and how may these underlying processes affect retention long after that information has left immediate awareness? These basic research questions have been among the utmost concern for research concerning the intersection of working memory (WM) and episodic long-term memory (LTM). WM refers to the immediate, capacity-limited memory system that underlies maintenance and manipulation of information for ongoing cognition, whereas LTM refers to the presumably unlimited retention of information that is no longer available in WM. There has been considerable interest in how purported processes may be involved in not only supporting maintenance in WM, but also later retrieval from episodic LTM (Camos & Portrat, 2015; Loaiza & McCabe, 2012; McCabe, 2008; Rose, Buchsbaum, & Craik, 2014).

A growing body of research has investigated refreshing as a candidate mechanism that supports maintenance in WM by directing attention to memoranda, and consequently, strengthening their activation in WM (see Camos et al., 2018 for a review). That is, by reactivating memory traces in the focus of attention, refreshing presumably augments the accessibility of the memoranda, particularly for their retrieval from WM. Although its functioning remains to be fully understood, thus far refreshing is generally considered to be an attention-based mechanism that has been distinguished both behaviorally and neurologically from other maintenance mechanisms, such as subvocal rehearsal (Valérie Camos, Lagner, & Barrouillet, 2009; Hudjetz & Oberauer, 2007; Loaiza & McCabe, 2013; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Vergauwe, Camos, & Barrouillet, 2014). One source of evidence for refreshing as an independent, attention-based mechanism has come from complex span tasks that are commonly used to measure WM capacity (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007). Whereas simple span tasks successively present memoranda (e.g., words), complex span tasks interleave the presentation of memoranda with secondary, distracting processing components (e.g., solving arithmetic problems) before their recall. Moreover, the relative attentional demand or cognitive load of the distractors can be manipulated (e.g., varying the pace of its presentation). These studies have collectively indicated a linear relationship between cognitive load and WM recall (Barrouillet, Portrat, & Camos, 2011), such that increasing the cognitive load of a distractor likewise reduces WM recall, irrespective of the match in modality between the memoranda and the distractors (i.e., verbal or visuospatial; Vergauwe, Barrouillet, & Camos, 2010) or the availability of rehearsal (Valérie Camos et al., 2009; Valérie Camos, Mora, & Barrouillet, 2013; Valérie Camos, Mora, & Oberauer, 2011; Mora & Camos, 2013). Such findings are commensurate with the notion that refreshing relies on a limited resource of domain-general attention. Other paradigms have explicitly instructed refreshing by presenting cues that prompt participants to "think of" the recently presented memoranda (Camos et al., 2011; Higgins & Johnson, 2009; Johnson, Reeder, Raye, & Mitchell, 2002; Souza & Oberauer, 2017a; Souza, Rerko, & Oberauer, 2015; Vergauwe & Langerock, 2017). This work has revealed that the frequency of refreshing directly improved visuospatial WM (Souza & Oberauer, 2017a; Souza et al., 2015), and that disruption of central rather than visual attention impaired visuospatial WM (Souza & Oberauer, 2017a). These findings collectively suggest that refreshing uses domain-general, central attention to bolster the activation of memoranda in WM.

If refreshing functions to strengthen the activation of memoranda in WM, it is reasonable to expect that these effects may endure long after information is no longer available in WM. Indeed, some findings have suggested that refreshing may also confer benefits for retrieval from episodic LTM (Valérie Camos & Portrat, 2015; M. K. Johnson et al., 2002). In our work (Loaiza & Borovanska, 2018; Loaiza, Duperreault, Rhodes, & McCabe, 2015; Loaiza & McCabe, 2012, 2013), we have investigated whether refreshing is the source of the *McCabe effect* (McCabe, 2008): the finding that, although initial recall is typically greater for memoranda studied during simple span compared to complex span, the reverse is true during final free recall, such that recall from complex span is more likely than simple span. Moreover, the difference between complex and simple span was greatest for words presented in earlier serial positions of the trial. McCabe (2008) originally asserted that the source of this effect is, compared to simple span, the relatively greater opportunity to covertly retrieve the memoranda back into immediate awareness during the pauses between the processing component and the next memorandum during complex span. This hypothesis drew on the theoretical models that presume that WM represents the capacity-limited subset of active representations in LTM (Cowan, 1999; Oberauer, 2002). That is, the central component of WM comprises a few privileged items, and this central component is embedded within activated LTM. The reasoning followed that due to the capacity limits of WM, memoranda are often displaced and must be retrieved back into the central component of WM in order to maintain their accessibility (Unsworth & Engle, 2007b, 2007a). Accordingly, the distraction during complex span tasks may displace the memoranda, thereby requiring their retrieval, whereas retrieval is not necessary for memoranda that have never been displaced, such as during simple span with only a few memoranda (Unsworth & Engle, 2006). Thus, although the distraction during complex span interferes with maintenance, it may also counterintuitively provide opportunities to retrieve the memoranda, with this practiced, internal retrieval benefitting episodic LTM the more it occurs. That is, if retrieval underlies the standard McCabe effect of greater long-term retention of for complex span compared to simple span items, then final free recall should increase as the number of distractors, and thereby retrieval opportunities, increase.

Furthermore, McCabe (2008) argued that covert retrieval is cumulative, such that each successively presented item has fewer opportunities to be retrieved in a typical complex span format that alternates memoranda with distraction. This is consistent with the finding that the McCabe effect is greatest for earlier serial positions than later ones, presumably due to more opportunities to retrieve earlier presented memoranda (Loaiza & Borovanska, 2018; Loaiza et al., 2015; McCabe, 2008). Indeed, when all of the memoranda were presented first followed by the distractors, the aforementioned negative slope across serial position was relatively flatter (Loaiza & McCabe, 2012). This suggested that adjusting the placement of distractors can vary the opportunity to cumulatively retrieve the memoranda in WM: all of the memoranda should have the same number of retrieval opportunities when the distractors follow their presentation, thereby yielding a relatively equivalent recall across serial position compared to alternating the distractors and memoranda in the typical complex span format.

Importantly, McCabe (2008) was agnostic regarding the precise nature of covert retrieval, postulating that it could “take the form of subvocal rehearsal (Baddeley, 1986), a simple mental search (Cowan, 1992), or a ‘refreshing’ process (Barrouillet et al., 2004)” (McCabe, 2008, p. 482). Given the finding that the McCabe effect was independent of the opportunity to engage in rehearsal (Loaiza & McCabe, 2013) and the strong similarity in concept between covert retrieval and refreshing, we had since argued that covert retrieval and refreshing are one and the same (Loaiza et al., 2015; Loaiza & McCabe, 2012, 2013; Loaiza, Rhodes, & Anglin, 2015). However, as we demonstrate in the following study, this conclusion may have been premature. Rather than refreshing still accessible memory traces via attention, the McCabe effect may reflect the effects of retrieving information from outside the central component of WM, much the same as the overt retrieval that occurs in many episodic LTM paradigms.

How can this be determined? One method would be to consider theoretically meaningful factors that may dissociate the two processes if they are indeed independent. For example, it is well known that increasing the number of presented memoranda, or list length (sometimes also referred to as *memory load* or *set size*), has systematic effects on memory performance. List length has been shown to reduce recall in standard immediate recall tasks (e.g., of sequentially presented words; Ward, Tan, & Grenfell-Essam, 2010), complex span tasks (Unsworth & Engle, 2006), short-term recognition tasks (Luck & Vogel, 1997), and delayed free recall tasks (Rohrer & Wixted, 1994). Furthermore, a classic finding in the memory literature is that the response speed to recognize whether a tested item was a member of a list of sequentially presented memoranda is directly related to list length (Sternberg, 1966, 1969). The results were interpreted as evidence of a search of recently presented information, with the search taking longer and thereby delaying response speed as list length increased. Such linear trends have also been shown for response speed during distractors of a complex span task (Jarrold, Tam, Baddeley, & Harvey, 2011; McCabe, 2010) and a Brown-Peterson span task (Vergauwe et al., 2014), with the latter study revealing that response latency increased about 40 ms with each additional item to maintain. This indicates that processing and storage may rely on a shared pool of attentional resources (Barrouillet & Camos, 2015), such that increasing list length incurs a cost that is evident when responding to another attentionally demanding activity. As detailed previously, given that refreshing is presumed to strongly rely on a limited resource of domain-general, central attention, these findings collectively suggest that list length may constrain refreshing if it is a mechanism that quickly cycles memoranda through the focus of attention in a serial manner (Vergauwe & Cowan, 2014, 2015). That is, the more items that are held in the central component of WM, the slower and more error-prone refreshing will be to serially reactivate their traces in the focus of attention.

Conversely, factors that demand attentional resources, such as list length, may have little impact on retrieving information from just outside the central component of WM (Oberauer, 2002, 2005). As discussed previously, embedded processes models distinguish between a few privileged items within the central component of WM and activated LTM that can comprise information activated above a baseline (Cowan, 1999; Oberauer, 2002, 2009). According to Oberauer's (2002, 2009) model, if this less available information becomes relevant, it is reloaded into the central component of WM (i.e., in Oberauer's model, the region of direct access) as a single chunk from outside the central component (i.e., activated LTM). The initial evidence for this notion of "chunking and unpacking" (Oberauer, 2009) has been demonstrated in asymmetrical list-switching costs: a once irrelevant list of items can become relevant once again and retrieved back into the central component of WM, but the number of items in the list has no effect on the efficiency of its retrieval (Oberauer, 2005). This indicates that the memory set is instead retrieved as a chunk rather than one-by-one into the focus of attention as is presumed to occur during refreshing. Other work comparing the time to access lists from WM compared to LTM has suggested that, although retrieval from episodic LTM takes longer overall than WM, list length has little additional effect (Conway & Engle, 1994; Wickens, Moody, & Dow, 1981). More recent neural evidence also supports this notion. In their study, Fukuda and Woodman (2017) measured suppression of alpha-band oscillations using electroencephalography (EEG) while participants learned and recalled the same four spatial arrays of variable set size (i.e., one array for each set size: 1, 2, 4, and 8 colored squares) that were interleaved multiple times amongst new arrays. When participants were later asked to retrieve one item of a set size 4 array from episodic LTM, the alpha-power suppression was initially similar to retrieving all four items, and then shifted to a pattern that resembled retrieval of just one item of a set size 1 array. This suggests that, at least up to four items, participants can retrieve the entire array from episodic LTM into WM.

At first glance, the notion that retrieval from outside the central component of WM can occur as a process of "chunking and unpacking" appears at odds with a variety of models from the extensive LTM literature that increasing list length reduces accuracy and increases recall latency (e.g., Rohrer & Wixted, 1994). Much of the literature emphasizes temporal-context (e.g., the list to which an item belongs or its serial position) as a pivotal cue for prompting retrieval from LTM (Howard & Kahana, 1999; Kahana, 1996; Unsworth, 2008). These temporal-contextual cues are thought to help to define a search set from which to sample items that depends on the list length and/or participants' individual abilities (e.g., Unsworth, 2007). However, these views are not necessarily incompatible: the "chunk" of items may be considered analogous to the contextual cue of list, which is first sampled and then used to retrieve items within the list as a two-stage search. Indeed, studies of episodic LTM (Howard, Youker, & Venkatadass, 2008; Unsworth, 2008) and semantic LTM (Gruenewald & Lockhead, 1980) do indicate that clustering of items based on their list context or membership of a category, respectively, can facilitate retrieval of individual items. Moreover, it is also important to note that the literature supporting the notion of the retrieval of chunks has focused on just a few items that are essentially at the standard limit of WM capacity (Cowan, 2000). Thus, it may be the case that the retrieval of a chunk or the cueing of a list may be less feasible if list lengths exceeded WM capacity.

Finally, in support of the notion that retrieval may be less attention-demanding than refreshing within the central component of WM, some work has indicated that dividing attention during retrieval has little effect on long-term retention compared to dividing attention during encoding (e.g., Buchin & Mulligan, 2017; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Murdock, 1965) or variation of cognitive load in complex span (Camos & Portrat, 2015; Jarjat et al., in press). Such results collectively suggest that factors that may impede upon a limited resource of central attention, such as cognitive load or list length, do not strongly affect retrieval of information that is no longer actively maintained in WM. Instead, such factors more strongly affect the active encoding and maintenance of those items within WM, with downstream consequences for their long-term retention.

Accordingly, the manipulation of list length would have important theoretical implications regarding whether retrieval and refreshing are redundant or distinguishable mechanisms. Figure 1 shows a schematic of the psychological processes and their respective expected pattern of results. If the benefit of distractors to episodic LTM is due to the refreshing of recently active memoranda one-by-one into the focus of attention (i.e., *the refreshing account*), then an interaction between list length and distractors would be expected: larger list lengths should attenuate the beneficial effect of distractors compared to smaller list lengths. That is, due to limited attentional resources, fewer memoranda would be able to be refreshed during trials with larger than smaller list lengths. Accordingly, it may become more difficult to refresh memoranda in WM as their number increases, thereby reducing the benefit of distractors for episodic LTM. Indeed, when examining the original serial position curves presented in McCabe (2008; Figure 3), one would expect that recall should be up to three times greater for serial position 1 of list length 4 than list length 2 due to three times as many distractors presented for the same serial position. However, as is common in complex span, list length was confounded with the number of distractors in this study. McCabe (2008; p. 490) noted that these findings may suggest that “memory load reduces the amount of covert retrieval that occurs…due to the attentional load imposed by already having to maintain items in memory.” As such, list length may attenuate the impact of distractors if the source of the benefit is refreshing. Conversely, if the source of the distractor benefit is due to retrieving information from outside of the central component of WM (i.e., *the retrieval account*), then the effect of distractors should not vary with list length given that the entire memory set is retrieved as a chunk. Such findings would provide a novel understanding of not only whether refreshing and retrieval are distinguishable, but also address the broader question of the overlap between WM and LTM. Namely, a distinction between refreshing and retrieval would suggest that refreshing is an attentional mechanism that operates on highly active information in WM and independently of LTM. Conversely, retrieval represents the recovery of information from outside of the central component of WM, and this can occur during LTM tasks as well as WM tasks. As such, retrieval during measures of WM may represent a way in which LTM processes influence WM (Loaiza, McCabe, Youngblood, Rose, & Myerson, 2011; Rose et al., 2014; Rose, Craik, & Buchsbaum, 2015; Rose, Myerson, Roediger, & Hale, 2010) and thereby blur the border between the systems (Loaiza & Camos, in press; Unsworth & Engle, 2007a).

The current experiments were designed to test these research questions. We presented Brown-Peterson span-like trials that manipulated the number of distractors (0 to 3) following a variable number of memoranda (2 to 4) to recall in serial order. After presentation of a block of trials and a short delay, participants freely recalled the words from the previous block. This study is also novel because increasing either the list length (Unsworth & Engle, 2006; Ward et al., 2010) or duration of a filled retention interval (Brown, 1958; Peterson & Peterson, 1959) are known to reduce recall, but these factors are rarely crossed in a WM paradigm. However, the primary interest was their impact on final free recall. We also considered final free recall conditionalized on correct initial recall to account for any potential impact of list length being simply an advantage of having recalled the smaller versus larger list lengths more successfully during initial recall. That is, we wanted to ensure that differences between conditions in episodic LTM were not merely due to any potential benefit of having differentially recalled the memoranda during presentation of the trials. It should be noted that in previous studies we had only considered unconditionalized performance (e.g., Loaiza & McCabe, 2012; McCabe, 2008). Nevertheless, given the importance of ensuring that a mere testing effect could not explain the pattern of results (e.g., a list length effect in final recall due to an increased likelihood of initially recalling smaller than larger list lengths), we included this measure as other researchers have done (e.g., Loaiza & Borovanska, 2018; Rose et al., 2014; Souza & Oberauer, 2017b).

Consistent with our previous work, we expected to find that increasing the number of distractors should improve final free recall despite being disadvantageous for initial recall (McCabe, 2008). As we have asserted previously, this may be due to the increased opportunities that the distractors afford to retrieve the memoranda in WM (Loaiza et al., 2015; Loaiza & McCabe, 2012). Moreover, because the distractors followed the memoranda, the retrieval opportunities should be constant across serial position. Accordingly, we predicted that final free recall should be relatively constant across serial position rather than negatively sloped as has been previously shown using a complex span format where the memoranda were alternated with the distractors (Loaiza & McCabe, 2012; McCabe, 2008).

The novel theoretical contribution of the experiments concerns how the effect of distractors may change with the amount of information to remember (i.e., list length). If refreshing underlies the beneficial effect of distractors to final free recall, then list length should interact with the number of distractors, such that their beneficial effect on final free recall is reduced with increasing list length. This should be particularly evident in an analysis of serial position: final free recall should be flatter across serial position for smaller than larger list lengths, indicating that increased list lengths reduce the opportunity for later-presented memoranda to be refreshed in the focus of attention compared to earlier-presented memoranda in the trial. However, if the benefit of distractors reflects retrieval of information outside of the central component of WM, then the advantage of increasing distractors for final free recall should occur regardless of list length. Moreover, final free recall should be flat across serial position regardless of list length, corroborating the notion that the memoranda are retrieved as a whole unit or chunk from activated LTM. Finally, since carrying out the first two experiments, a recent alternative account has suggested that the time to process information in WM, rather than refreshing or retrieval per se, underlies the advantage of the distractors for episodic LTM (i.e., *the total time account*; Jarjat et al., in press; Souza & Oberauer, 2017b). Thus, Experiment 3 was conducted to further address this possibility alongside the refreshing and retrieval accounts tested in Experiments 1 – 2.

Experiment 1

Method

Participants. Thirty-one participants (M_{age} = 22.50, *SD* = 2.78) were recruited through the department subject pool and compensated with £6 for the study that lasted approximately 60 min. Two participants were excluded from the analysis for not being native English speakers. Due to experiment failure, the first block of trials of one participant was lost and therefore excluded from analysis. All of the participants provided informed consent and were fully debriefed at the end for all of the experiments. The studies were approved by the Department of Psychology ethics committee at the University of Essex.

Design. The experiment manipulated the number of words to recall (2, 3, or 4) and the number of distractors following the words (0, 1, 2, or 3) within-subjects. We also examined recall performance as a function of serial position. The principal dependent variables were initial recall and final free recall (FFR).

Materials and Procedure. The memoranda for all of the experiments were randomly drawn from a pool of 154 concrete, high frequency nouns (letters: *M* = 5.35, *SD* = 1.29, range = 4 – 8; syllables: *M* = 1.47, *SD* = 0.50, range = 1 – 2; log HAL frequency: *M* = 9.29, *SD* = 0.96, range = 8.00 – 12.42) selected from the English Lexicon database (Balota et al., 2007). The words were randomly arranged for each participant. All experiments were programmed in Matlab with the Psychophysics toolbox extensions (Brainard, 1997; Kleiner et al., 2007).

During all of the experiments, an experimenter was present to individually test the participants, monitor their performance, and ensure compliance with the instructions.

Participants first practiced 10 arithmetic problems that would serve as the distractors during the task. The arithmetic problems (e.g., three + two = five?) were presented for 3.5 s with a 0.5 s interstimulus interval (ISI), and participants were instructed to read them aloud and respond true or false while pressing a corresponding key. Half were true and half were false. The participants remained in this practice phase until they reached an 85% criterion. Next, participants received instructions for the critical task. During the trials, a fixation point appeared on the screen for 1 s, followed by the first word for 1 s (0.5 s ISI) to read aloud and remember. Depending on the trial, another word followed for up to four words to recall. After all the words of the trial were presented, either a cue to recall the words immediately followed, or between one to three arithmetic problems were presented that were then followed by the cue to recall the words. The problems were read and solved aloud while pressing a corresponding key, as during the practice phase. When participants saw the cue to recall, they were instructed to try to recall the words aloud in the order they were originally presented (i.e., serial recall). The experimenter noted their recall on a score sheet after each trial and later coded the recorded speech files to enter initial recall for analysis. Participants completed three practice trials before the first block, and afterward, they received a summary of instructions for the critical trials and were further informed that their memory for the words would be tested again later in the experiment. These summarized instructions and warning about the future final recall test were repeated before each subsequent block of trials. There were four blocks, each comprising 12 trials (one trial per condition, e.g., 2 words to recall with 0 distractors) that were randomly presented for each participant.

After each block, participants completed a simple multiplication task for 1 min, followed by a final free recall test of the words that were presented in the previous block. Participants were instructed to freely recall the words without any regard to their order by typing them into the computer, and the recalled words appeared on the screen in a 6 x 6 grid (each block comprised 36 total words to recall). Participants were allowed to take a short break after each block. After all four blocks, participants completed an unrelated task and a short demographics questionnaire.

Results and Discussion

The results of both experiments were analyzed with Bayesian analysis of variance (BANOVA; Rouder, Morey, Speckman, & Province, 2012) with the independent variables (i.e., list length, distractors) as fixed effects and participant as a random effect using the BayesFactor package (Morey & Rouder, 2015) with its default settings in R (R core team, 2014). This method of analysis allows the comparison of the data given one model (e.g., the null model assuming only a random effect of participant, M_0) to that of another model (e.g., an alternative model assuming an effect of distractors, M_1). The ratio of these likelihoods is the Bayes factor (BF) that expresses the relative evidence for the alternative model (BF_{10}) or the null model (BF_{01}). One can also compare the relative evidence between models by examining the ratio between BFs associated with one model (e.g., a model including an effect of distractors) to that of another model (e.g., a model including effects of distractors and list length). The data and analysis scripts for all of the experiments can be found at http://osf.io/b9evq.

Initial Recall. Although serial recall was instructed, performance was also assessed without regard to the original serial position (i.e., free recall scoring). Both scoring measures were submitted to two separate 3 (list length: 2, 3, 4) x 4 (distractors: 0, 1, 2, 3) repeated measures BANOVAs (see Table 1). The best model for initial recall – serial scoring indicated strong effects and an interaction between list length and distractors ($BF_{10} = 2.54 \times 10^{47}$), but this model was not substantially preferred (BF = 2.26) to the simpler main effects only model ($BF_{10} = 1.12 \times 10^{47}$). For initial recall – free scoring, the main effects only model ($BF_{10} = 2.87 \times 10^{40}$) was best but not substantially preferred (BF = 1.98) to a model including an interaction between the factors (BF_{10} = 1.45 x 10^{40}). Overall, as list length and distractors increased, initial recall tended to decline, consistent with previous research (Unsworth & Engle, 2006; Ward et al., 2010).

For the sake of comparison to FFR, we also assessed initial recall performance as a function of serial position (see Figure 2A and 2B). For these analyses, performance for each list length was separately analyzed in three respective distractors x serial position repeated measures BANOVAs. These analyses indicated effects of distractors for all list lengths, but also effects of serial position for list lengths 3 and 4 in particular (serial scoring: BFs_{10} > 8.60 x 10^{35}; free scoring: BFs_{10} > 5.00 x 10^{17}). Thus, as the overall analysis demonstrated, initial recall tended to decline as the number of distractors increased, and further declined as a function of serial position particularly as list length increased. Such serial position effects have been commonly observed in list-learning paradigms (Ward et al., 2010).

Final free recall. FFR was the primary measure of interest, but as discussed previously, we also considered FFR conditionalized on initial recall to control for baseline differences across conditions (henceforth referred to as FFR*; e.g., Rose et al., 2014). Thus, FFR* refers to the later recall of memoranda that were initially recalled (approximately 78%, 73%, and 73% of the data in Experiments 1, 2, and 3, respectively). The few instances where values were undefined due to no initial recall were replaced with 0.

FFR and FFR* were submitted to two separate 3 (list length: 2, 3, 4) x 4 (distractors: 0, 1, 2, 3) repeated measures BANOVAs (see Figure 3). The FFR analysis showed no evidence of any effects of the factors (BFs_{10} < 0.80), whereas FFR* analysis showed that the best model included only a main effect of distractors (BF_{10} = 1,479) that was strongly preferred (BF = 31) to the next best model that included main effects of both distractors and list length (BF_{10} = 47). Thus, list length had little impact on final recall, and importantly, did not attenuate the beneficial effect of distractors on FFR*.

We further assessed final recall as a function of serial position to assess whether the expected relatively flat slope across serial position (Loaiza & McCabe, 2012) became increasingly negative as the list length increased. Accordingly, we submitted FFR and FFR* to three respective analyses for each list length as was performed for initial recall performance. In stark contrast to the analyses for initial recall, final recall was constant across serial position for each list length (see Figure 2C and Figure 2D). For FFR, there was no evidence of any effects of distractors or serial position ($BFs_{10} < 0.48$) except for list length 4, wherein a model of an effect of distractors ($BF_{10} = 449$) was strongly preferred (BF = 104) to a model including effects of distractors and serial position ($BF_{10} = 4$). For FFR*, the best models for each list length included an effect of distractors (BFs_{10} ranging from 10 – 4,970) that were preferred (BFs > 6) to the next best models including effects of distractors and serial position (BFs_{10} ranging from 1.31 – 108).

In summary, despite the unsurprising strong negative effect of list length on initial recall performance, list length did not affect FFR nor did it interact with the effect of distractors. Furthermore, inconsistent with our prior work (Loaiza &McCabe, 2012; McCabe, 2008), there was no evidence of a beneficial effect of the number of distractors on FFR. However, the strong evidence of an effect of distractors on FFR* and the descriptive statistics shown in Figure 3A indicate that diminished initial recall from trials with more distractors hindered the recall of those memoranda. The analysis of FFR* was included to ensure that the list length effect on initial recall did not attenuate the potential impact of list length on final recall. However, the analysis actually indicated that the number of distractors more strongly affected final recall when correcting for initial recall. Finally, final recall was constant across serial position, in line with the view that the beneficial effect of distractors reflects retrieval from outside the central component of WM rather than an attentional refreshing of memoranda within the focus of attention. For FFR*, the beneficial effect of the number of distractors occurred regardless of list length and the original serial position of the memoranda. If refreshing did underlie this effect, then list length should have attenuated the effect of distractors in final recall, indicating that the memoranda were increasingly more difficult to refresh as list length increased. At least in this study so far, the only moderator of the effect of the number of distractors appears to be having successfully recalled the memoranda in the first place.

It may be possible that these effects in final free recall were obfuscated by relatively low performance. Indeed, final free recall for memoranda with distractors in the current experiment ranged between 17-27% on average, whereas similar conditions in previous studies showed performance approaching and over 30% (Loaiza & McCabe, 2012; McCabe, 2008). There was also relatively high variability in the study, with the standard deviations sometimes approaching a similar value of the means of the conditions. Given these circumstances and the importance of replicating results in psychological science, we conducted Experiment 2 with a very similar design and set of goals as Experiment 1. The only change was a methodological one, such that the total number of words studied in each block was reduced in order to increase the overall likelihood of final free recall. Thus, the theoretical aims of Experiment 2 were the same as Experiment 1: assess whether list length moderates the advantage of increasing distractors on final free recall to establish a potential dissociation between refreshing and retrieval.

Experiment 2

Method

Participants. Thirty participants (M_{age} = 19.67, *SD* = 2.15) were recruited through the department subject pool and compensated with partial course credit. The study lasted approximately 60 min. None of the participants had participated in the previous experiment and all participants were native English speakers.

Design. The experiment manipulated the number of words to recall (2 or 4) and the number of distractors following the words (0, 1, 2, or 3) within-subjects. We also examined performance as a function of serial position. The principal dependent variables were initial recall and final free recall.

Materials and Procedure. The materials and procedure were very similar to Experiment 1, except that there were only 2 or 4 words to remember per trial. Thus, fewer total words were presented during each block compared to Experiment 1 in order to increase the final free recall performance. There were six blocks of eight trials each (one trial per condition, 24 total words per block) that were randomly presented for each participant. After finishing the critical task of the experiment, a portion of the participants completed an unrelated task and all participants completed a demographics questionnaire.

Results and Discussion

Initial recall. As in Experiment 1, free and serial initial recall scoring were submitted to two separate 2 (list length: 2, 4) x 4 (distractors: 0, 1, 2, 3) repeated measures BANOVAs. In general, these results replicated the findings from Experiment 1 and that of other research indicating strong negative effects of list length and distraction on immediate recall (see Table 1). The best model for serial scoring indicated effects of list length and distractors (BF_{10} = 2.57 x 10^{47}), and this model was not substantially preferred (BF = 1.92) to the model including an interaction term (BF_{10} = 1.34 x 10^{47}). For free scoring, a full model including an interaction term (BF_{10} = 3.32 x 10^{38}) was best but not substantially preferred (BF = 1.50) to the simpler main effects only model (BF_{10} = 2.22 x 10^{38}). The analyses regarding these effects as a function of serial position were similar to Experiment 1 (see Figure 4A and 4B): there was an effect of distractors at each list length, but also an effect of serial position for list length 4 (serial scoring: BF_{10} = 6.42 x 10^{79}; free scoring: BF_{10} = 7.06 x 10^{43}). Thus, increasing list length and number of distractors following the memoranda strongly reduced immediate recall, with performance increasingly declining as a function of serial position as list length increased.

Final Free Recall. FFR and FFR* were submitted to two separate 2 (list length: 2, 4) x 4 (distractors: 0, 1, 2, 3) repeated measures BANOVAs. Both measures indicated increased performance with increasing number of distractors (see Figure 5). For FFR, the best model included an effect of distractors (BF_{10} = 17) that was preferred (BF = 6) to the next best model that additionally included a main effect of list length (BF_{10} = 2.86). Surprisingly, the best model for FFR* showed beneficial effects of distractors *and* list length ($BF_{10} = 1.73 \times 10^9$) that was preferred (BF = 5) to the next best model of only an effect of distractors ($BF_{10} = 3.04 \times 10^8$). Finally, as in Experiment 1, we further considered final recall as a function of serial position. Once again, performance was constant across serial position (see Figure 4C and 4D): for FFR at list length 4, there was no evidence of any effects ($BFs_{10} < 0.21$), but the remaining analyses for FFR at list length 2 and FFR* showed consistent evidence for an effect of distractors (BFs_{10} ranging from 715 – 4.86×10^6) that was preferred (BFs ranging from 5 – 118) to the next best model that further included an effect of serial position (BFs_{10} ranging from 132 – 7.29×10^5).

In summary, the results of Experiment 2 partially diverged from those of Experiment 1 in showing more consistent evidence of a beneficial effect of distractors on final free recall. As in Experiment 1, this effect was constant across serial position and larger when correcting for initial recall, but a beneficial effect of list length also emerged alongside distractors in this conditionalized analysis. Given this unexpected finding, we conducted a further exploratory analysis to consider the source of this effect. Namely, it may have been the case that recalling a word from a trial prompted the recall of further words from the same trial. Given that larger list lengths necessarily comprised more candidate words to recall than smaller list lengths, the likelihood of recalling from the same trial may have asymmetrically benefitted larger list lengths. Accordingly, we examined the likelihood of successively recalling memoranda from the same trial during final free recall in a 2 (list length: 2, 4) x 4 (distractors: 0, 1, 2, 3) repeated measures BANOVA. Consistent with this notion, the best model included main effects of list length and distractors (BF_{10} = 1.84 x 10^5), and this model was preferred (BF = 7) to a model that included only an effect of distractors (BF_{10} = 23,603). Thus, just as retrieval may promote the likelihood of using the original temporal-contextual cues from the trial to guide retrieval from episodic LTM (Loaiza & McCabe, 2012), perhaps the associations between the memoranda of larger list lengths are also encoded and utilized during retrieval from episodic LTM. Further work will be necessary to explore this possibility. Notwithstanding this result, the findings converge with the previous experiment in suggesting that the presentation of distractors facilitates later retrieval from episodic LTM. Importantly, the consistency of the effect across list length and serial position indicates that the effect is more likely due to retrieval from outside of the central component of WM rather than refreshing the memoranda through the focus of attention.

Although the results are consistent with the retrieval account, they could also be interpreted as consistent with a recent alternative account suggesting that prolonging the time that memoranda are processed in WM facilitates their retrieval from episodic LTM (Jarjat et al., in press; Souza & Oberauer, 2017b). Thus, it is time, rather than the distractors per se, that improves long-term retention. In a recent series of experiments, Souza and Oberauer’s (2017b) results indicated that a slow span trial that presented a blank pause between the memoranda of an equivalent duration to that of the distracters during complex span yielded greater final free recall compared to the complex span and simple span trials. The slow span trials further showed greater final free recall compared to trials in which the distractors or the blank pauses followed the presentation of all of the memoranda in a Brown-Peterson-like format. The authors concluded prolonging the time that memoranda are attended to just after they have been encoded, perhaps by consolidation (e.g., Bayliss, Bogdanovs, & Jarrold, 2015) or elaboration (e.g., Craik & Tulving, 1975), improves their retention in episodic LTM. Consolidation refers to transforming sensory input into a stable short-term trace, whereas elaboration refers to processing items with regard to their meaningful, semantic characteristics.

Accordingly, this account seems consistent with the pattern of results here: increasing number of distractors following the memoranda in the current experiments necessarily increased the total duration that they were processed in WM; time and distractors in these experiments are confounded. Accordingly, the beneficial effect of increasing distractors on episodic LTM may be due to the increased time rather than retrieval from outside the central component of WM. One inconsistency with this alternative account in the current results, however, is that a benefit of distractors was observed when they followed all of the memoranda. The beneficial effect of the distractors may not have been expected at all in the current study given that the benefit of time in Souza and Oberauer’s study appeared to be exclusive to the slow span trials wherein the pauses followed each item rather than all of the items. Thus, it is at least unlikely that the memoranda in the current study benefitted from improved consolidation given that the increased time did not follow each item. However, it may still have been the case that the increased time afforded greater elaboration or refreshing during the free time available during the task (Jarjat et al., in press).

Experiment 3 was designed to distinguish these accounts while remaining within the scope of the original goal of the study to examine the potential impact of list length and its interaction with distractors on final free recall. Namely, we again varied the number of memoranda (2 or 4) presented during the trial, but the number of distractors changed slightly from the previous experiments. Either 0 or 3 distractors followed the memoranda, but the 3 distractors were either presented at the rate from the previous experiments (3.5 s) or at a slower rate (4.5 s). This allows a disambiguation between the number of distractors and the total time that the items are processed in WM. Although both the retrieval and free time accounts would predict a benefit of 3 over 0 distractors in final free recall, they differ regarding the potential difference between the fast and slow presentation rates of the 3 distractors. If the time afforded to the items to be processed in WM is the true cause of the previously reported benefit of increasing distractors, then final free recall should be even greater for the memoranda studied during the 3 slow distractors (15 s total before recall) relative to the 3 fast distractors (12 s total before recall). By contrast, the retrieval account suggests the number of distractors affording retrieval opportunities primarily facilitates final free recall in this paradigm, and thus there should be no difference in final free recall between the fast and slow presentation rates of the 3 distractors.

Experiment 3

Method

Participants. Thirty participants (M_{age} = 19.83, *SD* = 2.28) were recruited through the department subject pool and compensated with partial course credit or £5 for the study that lasted approximately 45 min. None of the participants had participated in the previous experiment and all participants were native English speakers.

Design. The experiment manipulated the number of words to recall (2 or 4) and the number of distractors following the words (0, 3 distractors fast, 3 distractors slow) within-subjects. We also examined performance as a function of serial position. The principal dependent variables were initial recall and final free recall.

Materials and Procedure. The materials and procedure were very similar to the previous experiments. The main differences were that only 0 or 3 distractors followed the 2 or 4 memoranda. Importantly, the presentation rate of the 3 distractors varied across the trials: they were presented either for 3.5 s (0.5 s ISI) as in the previous experiments or for a slower rate of 4.5 s (0.5 s ISI). Participants were not alerted to this difference. There were six blocks of six trials each (one trial per condition, 18 total words per block) that were randomly presented for each participant. The participants finished the session at the conclusion of the last block, with no further tasks completed.

Results and Discussion

Initial Recall. As in the previous experiments, serial and free recall scoring were both assessed as measures of initial recall performance (see Table 1). The 2 (list length: 2, 4) x 3 (distractors: 0, 3 fast, 3 slow) repeated measures BANOVA for both measures showed that the best model was a full model including an interaction between list length and distractors (serial scoring: BF_{10} = 1.53 x 10^{50}; free scoring: BF_{10} = 9.75 x 10^{38}), and these models were substantially preferred (serial scoring: BF = 1,645, free scoring: BF = 40) to the next best models including only main effects (serial scoring: BF_{10} = 9.31 x 10^{46}; free scoring: BF_{10} = 2.39 x 10^{37}). The source of the interaction was a relatively greater difference between list lengths for the 3 distractors conditions compared to 0 distractors condition. The analyses regarding these effects as a function of serial position were similar to the previous experiments (see Figure 6A and 6B): there were clear effects of distractors at each list length, with more substantial evidence for effects of serial position at list length 4 (serial scoring: BF_{10} = 4.51 x 10^{76}; free scoring: BF_{10} = 2.19 x 10^{48}). Thus, the results are consistent with the previous experiments and previous work showing the strong impairments of increasing list length and distraction to recall, with declines also evident as a function of serial position.

Final Free Recall. As in the previous experiments, FFR and FFR* were submitted to two separate 2 (list length: 2, 4) x 3 (distractors: 0, 3 fast, 3 slow) repeated measures BANOVAs. Once again, both measures indicated a benefit of distractors, such that final recall was better for 3 distractors over 0 distractors (see Figure 7). For FFR, the best model included an effect of distractors (BF_{10} = 7.29) that was preferred (BF = 6) to the next best model that additionally included a main effect of list length ($BF_{10} = 1.21$). For FFR*, the best model was the full model including an interaction between distractors and list length ($BF_{10} = 7.57 \times 10^8$), but this model was not substantially preferred (BF = 1.35) to the simpler model including only main effects ($BF_{10} = 5.61 \times 10^8$). The source of the interaction appeared to originate from a greater effect of distractors for list length 4 than list length 2. It is not immediately clear how such a result would comport with either the retrieval or free time accounts. Given the results of Experiment 2, it was possible that increasing list length could improve final free recall overall perhaps due to the relatively increased temporal associations that are possible between items compared to small list lengths. Notwithstanding, the most important analysis to adjudicate between the retrieval and free time accounts concerned the difference between the 3 distractors conditions. Thus, our analysis primarily concerned this pairwise comparison between the conditions and their respective differences with the 0 distractors condition.

In order to assess the differences between these conditions, we conducted directional, one-way Bayesian *t*-tests (Rouder, Speckman, Sun, Morey, & Iverson, 2009) that assumed an advantage of the 3 distractors conditions over the 0 distractors condition as well as non-directional two-way Bayesian *t*-tests between the 3 distractors conditions. In addition, we also used Bayesian estimation software (BEST; Kruschke, 2013) to estimate the size of the effects for each comparison and their respective 95% highest-density intervals (HDI). For FFR, we performed this analysis on the average of the two list lengths given the lack of list length effect, whereas the conditions were compared for each respective list length for FFR* given the evidence of a list length effect.

As both retrieval and free time accounts would predict, a McCabe effect was evident in FFR, such that both of the 3 distractors conditions showed an advantage over the 0 distractors condition (fast: $BF_{10} = 16$; slow: $BF_{10} = 13$). Importantly, consistent with the retrieval account, there was substantial evidence for a null difference between the two 3 distractors conditions (BF_{10} = 0.22). The effect sizes for these comparisons derived with BEST were convergent with these findings: there were moderate benefits of the 3 distractors conditions over the 0 distractors condition (fast: *d* = 0.56 [0.15, 0.97]; slow: *d* = 0.58 [0.15, 1.02]). However, there was no evidence of a difference between the two 3 distractors conditions, with the range of credible values of the effect size including 0 (*d* = 0.09 [-0.29, 0.45]). This pattern was similar for the analysis of FFR*: there was a McCabe effect when comparing the 0 distractors condition to both of the 3 distractors conditions for list length 2 (fast: BF_{10} = 64, *d* = 0.66 [0.25, 1.01]; slow: BF_{10} = 5, *d* = 0.49 [0.07, 0.92]) and list length 4 (fast: BF_{10} = 68,047, *d* = 1.17 [0.67, 1.66]; slow: BF_{10} = 66,673, *d* = 1.27 [0.66, 1.87]). The comparison between the two 3 distractors conditions, however, supported a null effect (list length 2: BF_{10} = 0.24, *d* = -0.12 [-0.50, 0.25]; list length 4: BF_{10} = 0.56, *d* = 0.28 [-0.10, 0.67]).

Finally, as in the previous experiments, we also examined final recall as a function of serial position. Congruent with the previous findings, performance was consistent across serial position (see Figure 6C and 6D): for FFR at list length 2, there was no evidence of any effects (BFs_{10} < 1.90), but the remaining analyses for FFR at list length 4 and FFR* showed consistent evidence for an effect of distractors (BFs_{10} ranging from 111 – 1.09 x 10^{15}) that was preferred (BFs ranging from 3 – 30) to the next best model that further included an effect of serial position (BFs_{10} ranging from 4 – 3.62 x 10^{13}).

Thus, the results of Experiment 3 replicate and extend those of the previous experiments to show that the benefit of distraction to final free recall is more likely due to the impact of retrieval during WM rather than the amount of time to process the memoranda in WM per se. Moreover, they are consistent with the previous experiments in that list length did not moderate the beneficial effect of distractors, supporting the notion that retrieval serves to recover information from outside of the central component of WM.

General Discussion

A classic finding in the memory literature is that performance tends to decline as the duration of a distraction-filled retention interval increases (Brown, 1958; Peterson & Peterson, 1959). Counterintuitive to this result, the overall findings of this study suggest that increasing the number of distracting activities in a WM task promotes later retrieval from episodic LTM. This is consistent with prior work showing an advantage of complex span over simple span items for retrieval from episodic LTM, i.e., the McCabe effect (Loaiza et al., 2015; Loaiza & McCabe, 2012, 2013; McCabe, 2008; Souza & Oberauer, 2017b). We have interpreted such findings as consistent with a covert retrieval account of WM (McCabe, 2008): although distractors during span tasks interfere with maintenance in WM, they also serve as opportunities to covertly retrieve recently presented memoranda, in turn reinforcing the contextual cues that can guide later retrieval from episodic LTM (Loaiza & McCabe, 2012; McCabe, 2008). The current study replicated and extended these results by demonstrating that the number of memoranda to recall, or list length, did not moderate this benefit, and the benefit was consistent across serial position. Furthermore, the results were able to rule out an alternative competitive account that extending the time with which to process items in WM primarily underlies the beneficial effect of distraction to episodic LTM (Jarjat et al., in press; Souza & Oberauer, 2017b). Overall, the study makes a novel contribution to the literature to suggest that retrieval may be dissociable from refreshing, especially regarding its impact on episodic LTM. Such findings also have implications for the overlap between WM and LTM.

The Distinction between Refreshing and Retrieval in WM

The novel manipulation of list length in the current study was intended to dissociate two possible underlying causes of a counterintuitive result that distractors can facilitate episodic LTM (McCabe, 2008): (1) the memoranda are refreshed in the focus of attention (Loaiza et al., 2015; Loaiza & McCabe, 2012), or (2) retrieval of a cohesive chunk from outside of the central component of WM, and as such, retrieval and refreshing are dissociable (Camos et al., 2018; Loaiza & Camos, 2017). These possible conceptualizations are represented in Figure 1. As McCabe (2008) suggested, increasing list length may diminish the refreshing of memoranda one-by-one within the focus of attention. That is, given that increasing list length may incur an attentional cost (e.g., Vergauwe et al., 2014), refreshing may be slower and more error-prone as list length increases if it serially cycles through each item, thereby reducing the benefit of the opportunity to refresh the memoranda after each distractor. Accordingly, a reduced beneficial effect of distractors with increasing list length in the current study would be consistent with the refreshing account. Furthermore, examining final free recall as a function of serial position would provide further evidence for this view. Specifically, a relatively consistent pattern of final free recall across serial position was demonstrated using a Brown-Peterson span-like trial relative to the negative recency effect evident in typical complex span (Loaiza & McCabe, 2012). Consistent with a serial, cumulative view of refreshing, this predicted flat slope of final free recall across serial position may be more evident for small rather than large list lengths, indicating that later-presented memoranda in the trial have less opportunity to be refreshed than earlier-presented memoranda. In summary, the refreshing account would predict an interaction between list length and distractors, such that the beneficial effect of distractors is reduced for larger list lengths and items toward the end of a trial.

In stark contrast to these predictions, list length never moderated the beneficial effect of distractors on final free recall across all three experiments. Furthermore, we replicated and extended Loaiza and McCabe's (2012) finding that final free recall is consistent across serial position for Brown-Peterson span-like trials regardless of list length. The results instead comport with the retrieval account, such that retrieving the entire chunk of items from outside the central component of WM facilitates episodic LTM. In this regard, retrieval is less attention-demanding compared to refreshing, and may share more in common with the overt retrieval that occurs during many episodic LTM paradigms. We return to this potential overlap between WM and LTM in the next section.

These results are at first seemingly in conflict with the aforementioned observation that list length may have diminished the efficiency of retrieval in McCabe's (2008) original study. Instead, this may have been due to the increased demands of the distractors that are often confounded with list length in complex span tasks. That is, complex span tasks typically present one distractor for each memorandum (Conway et al., 2005). The current study sheds light on this by de-confounding list length and distractors within the same task. The results previously indicating a moderating effect of list length may instead be more attributable to the consequences of managing the demands of increasingly more distractors.

Most importantly, list length provided a theoretically meaningful factor along which the distinction between refreshing and retrieval was possible. As discussed previously, list length effects in WM have often been interpreted as reflecting the inefficiency of the system to actively maintain too many memoranda at once (e.g., Cowan, 2001; Oberauer, 2002). For example, list length effects are commonly shown in both accuracy (e.g., Luck & Vogel, 1997) and response speed (e.g., Oberauer, 2002) for items that are actively maintained in WM. A growing body of research suggests that refreshing is a domain-general mechanism that relies on central attention to boost the activation of items by cycling them through the focus of attention (see Camos et al., 2018 for a review). As such, a source of list length impairments to WM may originate from a reduced efficiency of refreshing to search and reactivate items in WM, suggesting a strong link between list length and attention. By contrast, list length may have less effect on retrieving a chunk or list from outside the central component of WM, such as recently presented but less available information in activated LTM (e.g., Oberauer, 2005) or information in episodic LTM (e.g., Fukuda & Woodman, 2017). Accordingly, the impact of list length and its interaction with distractors in the current study provides a means of delineating refreshing within or retrieval outside of the central component of WM.

Although in previous research we have made the argument that the McCabe effect reflects refreshing (e.g., Loaiza et al., 2015; Loaiza & McCabe, 2012, 2013), these findings indicate that refreshing and retrieval are in fact dissociable (Jarjat et al., in press; Loaiza & Camos, 2018). We thus contend that the dissociation that has been made in the literature regarding attentional refreshing and articulatory rehearsal (e.g., Camos et al., 2009) may be analogous to the distinction we make here: rehearsal, refreshing, and retrieval may operate independently to sustain memory traces in WM and, as a function of their use, have differential effects on the long-term retention of those traces. Refreshing and retrieval are more similar to each other in this regard in that they both have been shown to impact episodic LTM compared to rehearsal (e.g., Camos & Portrat, 2015; Loaiza & McCabe, 2013). On the other hand, rehearsal and retrieval may be more similar to each other in that they incur little attentional demand compared to refreshing (e.g., Vergauwe et al., 2014). These mechanisms may be differentially emphasized or strategically used as a function of the characteristics of the task or individual, as has been suggested regarding the joint use of rehearsal and refreshing (Camos et al., 2011). The contention that rehearsal, refreshing, and retrieval are dissociable mechanisms is consistent with the Multiple-Entry, Modular Memory System (MEM) framework (M. K. Johnson, 1992) that originally advanced the notion that these and other different reflective mental operations contribute to WM and episodic LTM. Moreover, Oberauer's (2002) concentric framework provides an architecture that could feasibly map onto the operations of refreshing and retrieval, such that refreshing cycles through a few active memoranda within the region of direct access, whereas retrieval recovers less active information that has been relegated to activated LTM. As we discuss further in the next section, this dissociation between refreshing and retrieval may further shed light on what previously appeared to be incompatible findings, such as whether refreshing relies on LTM (Loaiza et al., 2015; Rose et al., 2014).

The current results are also relevant to recent work examining the effects of varying cognitive load on WM and episodic LTM alike (Camos & Portrat, 2015; Jarjat et al., in press). As discussed previously, much of the work regarding cognitive load has conceptualized this factor as fundamentally linked to refreshing: as the cognitive load of presented distractors increase, the ability to engage in refreshing is reduced, thereby leading to reduced recall from WM (e.g., Barrouillet et al., 2011). A common method of manipulating cognitive load is to vary the presentation rate (i.e., *pace*) of the distractors, such that a short pace is thought to incur a high cognitive load relative to a long pace. In Experiment 3, the pace of the distractors was manipulated, and so at first glance the results may seem in conflict with these recent studies that have shown that manipulating refreshing via the pace of the distractors affects episodic LTM (Camos & Portrat, 2015; Jarjat et al., in press).

There are two reasons why the current results should not be interpreted as such. First, our experiments used a Brown-Peterson task, and slowing the pace necessarily increases the retention interval between study and test. Increasing retention intervals during a Brown-Peterson task are traditionally associated with profound forgetting (Brown, 1958; Peterson & Peterson, 1959). This contrasts with the impact of slowing the pace of distractors between items in complex span, such that a slow pace is thought to reduce cognitive load and thus reduce forgetting in WM, as discussed previously. The fact that there is no impact of the pace of the distractors on initial recall in Experiment 3 perhaps suggests that the benefit of reduced cognitive load was balanced against the cost of an increasing retention interval. Notwithstanding, Experiment 3 was not intended to be a test of cognitive load on episodic LTM. If this were the goal, a more appropriate method would be to manipulate difficulty while keeping pace constant as some others have done when examining initial recall (Ricker, Vergauwe, Hinrichs, Blume, & Cowan, 2015; White, 2012). Thus, these factors in combination with the fact that no pace effect was observed in initial recall suggests that further interpretation about cognitive load in this paradigm should be approached with caution. Secondly, as discussed previously, different tasks and manipulations may promote different maintenance mechanisms. For example, Vergauwe et al. (in press) showed that whether consolidation or refreshing is preferred as a strategy appears to depend on the temporal parameters of the task. Similarly, utilizing an arithmetic problem as a distractor over an extended period of several seconds may more strongly emphasize retrieval than refreshing compared to a paradigm that emphasizes quick, successive distractors, such as digits (Camos & Portrat, 2015; Jarjat et al., in press). These are certainly issues that require future research, but as we have reiterated previously, it is likely that a variety of different strategies (e.g., rehearsal, refreshing, elaboration, etc.) are used to maintain items in WM, and a growing body of work has sought to distinguish them (e.g., Camos et al., 2011; Vergauwe et al., in press). The current research contributes to this area by showing that there may be a dissociation between two of these mechanisms, namely, refreshing and retrieval.

Retrieval and the Overlap between Working Memory and Long-Term Memory

The distinction between WM and LTM has often been a considerable source of debate in the memory literature, with some espousing complete independence between the systems (e.g., Baddeley, 1986; Barrouillet & Camos, 2015) whereas others view memory as unitary with no need for a distinction between WM and LTM (e.g., Crowder, 1982; Nairne, 2002). An intermediate view suggests that WM represents the activated content of LTM (e.g., Cowan, 1999; Oberauer, 2002). At the very least, WM and LTM are often shown to be strongly related constructs, such that performance on a number of WM and LTM tasks correlate at the latent level (Unsworth, 2010). Furthermore, the growing literature regarding the impact of WM mechanisms on episodic LTM (Camos & Portrat, 2015; Loaiza & McCabe, 2012, 2013; McCabe, 2008) and likewise the effect of LTM processes on WM functioning (Loaiza & Camos, in press; Loaiza et al., 2011; Rose et al., 2014) point to a bidirectional relationship between the two systems. The present study suggests that retrieval in WM may be a means in which LTM processes influence WM, thereby blurring the border between the two systems.

In addition to the current study suggesting that list length has no impact on the efficacy of retrieval, there is further evidence in the literature that this is the case. For example, Rose and colleagues (2014, 2015) have shown a variable levels-of-processing effect in WM depending on the demands of a concurrent task. The levels-of-processing effect refers to the memorial advantage of processing information with regard to its deep, semantic characteristics compared to its shallow characteristics (Craik & Tulving, 1975). In Rose and colleagues' studies, participants were instructed to make deep or shallow processing decisions on memoranda before performing a concurrent task: either rehearsing the memoranda, solving simple or difficult arithmetic problems. These conditions were thought to vary the use of rehearsal and refreshing, with both being largely unavailable in the difficult condition. Rose and colleagues (2014, 2015) demonstrated an increasing levels-of-processing effect in WM as the difficulty of the concurrent task increased. To account for these and other findings, Rose and colleagues (Rose et al., 2014, 2015, 2010; Rose & Craik, 2012) have advanced a dynamic processing model of WM that episodic-retrieval processes influence maintenance in WM in addition to other mechanisms like rehearsal. That is, when the demands of a concurrent task deter the use of rehearsal and refreshing, retrieval from episodic LTM will help to sustain short-term retention of items in WM. The findings of this study are in line with this proposal, such that retrieval operates to recover information from outside the central component of WM.

The notion that retrieval represents an intersection between WM and LTM also helps to elucidate previously incongruent findings regarding whether refreshing relies on LTM. As Camos and colleagues' (2018) review notes, refreshing is fundamentally an attention-based mechanism and as such should hypothetically have little reliance on LTM, although research has not always been consistent regarding this issue (Ricker & Cowan, 2010). Loaiza and colleagues (2015) showed that the McCabe effect was not evident for non-words that likewise have no existing representation in LTM. This may make sense considering the results of the present findings: if retrieval is primarily responsible for the McCabe effect and we have now demonstrated that it may be more akin to retrieval that occurs during episodic LTM paradigms, then a moderating effect of representations' existing status in LTM would certainly be expected. In this regard, retrieval may rely more on LTM characteristics and less on attention, whereas refreshing may rely more on attention and less on LTM characteristics.

Limitations

There are several conceptual and methodological limitations to consider for this series of experiments. First, and most importantly, the concept of refreshing is still not fully understood, and some researchers disagree on fundamental aspects of its functioning (Camos et al., 2018). Most relevantly, the timing of refreshing is debated, with some work has suggested that refreshing occurs very rapidly, within about 40 ms (Vergauwe et al., 2014), whereas other work suggests that the time course of refreshing is much slower, about 600 ms (Johnson, McCarthy, Muller, Brudner, & Johnson, 2015). Still others argue that refreshing may even operate in parallel rather than in a serial manner (Portrat & Lemaire, 2015). Thus, the predictions regarding the interaction of list length and distractors in the work presented here very much rely on a specific depiction of refreshing as a relatively slow and serial mechanism. If refreshing could operate on multiple items in parallel or rapidly cycle through representations in the focus of attention, then the time allotted during the distractors in the current study likely could have been sufficient to refresh the items multiple times, thereby yielding a null interaction with list length. Unfortunately, because the paradigm is necessarily inferential, and no strategy reports were systematically recorded, it is difficult to discern between these possibilities. This limitation of the paradigm's indirectness also limits how much other strategies, such as elaboration or association between items, contributed to the results.

Furthermore, the argument that retrieval could occur as a chunk has largely originated from studies that utilized response times as the measure of interest rather than recall accuracy used here (Conway & Engle, 1994; Oberauer, 2005; Wickens et al., 1981). It may be the case that accuracy is differently sensitive to list length effects than response times, and thus the measures may not necessarily be congruent. Furthermore, list length effects are commonly observed in accuracy in episodic LTM paradigms (e.g., Rohrer & Wixted, 1994; Unsworth, 2007), thus calling the notion of retrieval of chunks into question. However, as we discussed previously, this literature is not necessarily at odds with the suggestion given that a cue-dependent search of LTM often involves retrieving a list context from which to sample specific items (Howard et al., 2008; Unsworth, 2008). That said, further work is necessary to ensure that these views are empirically as well as conceptually compatible.

Finally, a methodological issue that is persistent across this and other studies that have investigated the impact of WM mechanisms on episodic LTM is the confounding of retrieval instructions with time of test. Traditionally, participants are instructed to recall items in their original serial order for measures of WM, whereas such an instruction would be very difficult to accomplish in a test of episodic LTM. Thus, more work should be done to ensure that any of the results that have been observed in this area are not dependent on the difference in retrieval instructions between the two time points of recall.

Conclusions

The current study demonstrates novel findings regarding the interaction between list length and distractors presented during a span task trial. In particular, we suggest that the lack of a moderating effect of list length on the beneficial effect of distractors to final free recall indicates that retrieval may not strongly rely on attention and is therefore distinguishable from refreshing. Thus, we rescind our previous proposal that the McCabe effect may reflect refreshing, and instead suggest that refreshing and retrieval are two mechanisms that jointly operate in WM, with retrieval being a means by which LTM factors may influence WM.

References

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Barrouillet, P., & Camos, V. (2015). *Working Memory: Loss and reconstruction* (1 edition). East Sussex: Psychology Press.

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, *118*(2), 175–192. https://doi.org/10.1037/a0022324

Bayliss, D. M., Bogdanovs, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language, 81,* 34–50. https://doi.org/10.1016/j.jml.2014.12.004

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897x00357

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, *10*(1), 12–21. https://doi.org/10.1080/17470215808416249

Buchin, Z. L., & Mulligan, N. W. (2017). The Testing Effect Under Divided Attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000427

Camos, Valerie, Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *0*(0). https://doi.org/10.1111/nyas.13616

Camos, Valérie, Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*(3), 457–469. https://doi.org/10.1016/j.jml.2009.06.002

Camos, Valérie, Mora, G., & Barrouillet, P. (2013). Phonological similarity effect in complex span task. *The Quarterly Journal of Experimental Psychology*, *66*(10), 1927–1950. https://doi.org/10.1080/17470218.2013.768275

Camos, Valérie, Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, *39*(2), 231–244. https://doi.org/10.3758/s13421-010-0011-x

Camos, Valérie, & Portrat, S. (2015). The impact of cognitive load on delayed recall. *Psychonomic Bulletin & Review*, *22*(4), 1029–1034. https://doi.org/10.3758/s13423-014-0772-5

Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786.

Cousineau, D. (2005). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*, 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, *31*, 668–684.

Cowan, N. (1999). An embedded process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 62–101). New York: Cambridge University Press.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185. https://doi.org/10.1017/S0140525x01003922

Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology. General*, *125*(2), 159–180.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Crowder, R. G. (1982). The demise of short-term-memory. *Acta Psychologica*, *50*(3), 291–323. https://doi.org/10.1016/0001-6918(82)90044-0

Fukuda, K., & Woodman, G. F. (2017). Visual working memory buffers information retrieved from visual long-term memory. *Proceedings of the National Academy of Sciences*, *114*(20), 5306–5311. https://doi.org/10.1073/pnas.1617874114

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology Human Learning & Memory*, *6*(3), 225–240.

Higgins, J. A., & Johnson, M. K. (2009). The consequence of refreshing for access to nonselected items in young and older adults. *Memory & Cognition*, *37*(2), 164–174. https://doi.org/10.3758/MC.37.2.164

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(4), 923–941.

Howard, Marc W., Youker, T. E., & Venkatadass, V. S. (2008). The persistence of memory: Contiguity effects across hundreds of seconds. *Psychonomic Bulletin & Review*, *15*(1), 58–63. https://doi.org/10.3758/PBR.15.1.58

Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory & Cognition*, *35*(7), 1675–1684.

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

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2011). How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 688–705. https://doi.org/10.1037/a0022527

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

Johnson, M. K., Reeder, J. A., Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychological Science, 13*(1), 64–67. https://doi.org/10.1111/1467-9280.00411

Johnson, M. R., McCarthy, G., Muller, K. A., Brudner, S. N., & Johnson, M. K. (2015). Electrophysiological Correlates of Refreshing: Event-related Potentials Associated with Directing Reflective Attention to Face, Scene, or Word Representations. *Journal of Cognitive Neuroscience*, *27*(9), 1823–1839. https://doi.org/10.1162/jocn_a_00823

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & others. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. https://doi.org/10.1037/a0029146

Loaiza, V. M., & Borovanska, B. M. (2018). Covert retrieval in working memory impacts the phenomenological characteristics remembered during episodic memory. *Consciousness and Cognition*, *57*(Supplement C), 20–32. https://doi.org/10.1016/j.concog.2017.11.002

Loaiza, V. M., & Camos, V. (in press). The role of semantic representations in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Loaiza, V. M., & Camos, V. (2018). The long-term consequences of attentional refreshing and covert retrieval in working memory. *Manuscript under Review*.

Loaiza, V. M., Duperreault, K. A., Rhodes, M. G., & McCabe, D. P. (2015). Long-term semantic representations moderate the effect of attentional refreshing on episodic memory. *Psychonomic Bulletin & Review*, *22*(1), 274–280. https://doi.org/10.3758/s13423-014-0673-7

Loaiza, V. M., & McCabe, D. P. (2012). Temporal–contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Memory & Cognition*, *40*(2), 191–203. https://doi.org/10.3758/s13421-011-0148-2

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

Loaiza, V. M., McCabe, D. P., Youngblood, J. L., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1258–1263. https://doi.org/10.1037/a0023923

Loaiza, V. M., Rhodes, M. G., & Anglin, J. (2015). The Influence of Age-Related Differences in Prior Knowledge and Attentional Refreshing Opportunities on Episodic Memory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *70*(5), 729–736. https://doi.org/10.1093/geronb/gbt119

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, *58*(2), 480–494. https://doi.org/10.1016/j.jml.2007.04.004

McCabe, D. P. (2010). The influence of complex working memory span task administration methods on prediction of higher level cognition and metacognitive control of response times. *Memory & Cognition*, *38*(7), 868–882. https://doi.org/10.3758/MC.38.7.868

Mora, G., & Camos, V. (2013). Two Systems of Maintenance in Verbal Working Memory: Evidence from the Word Length Effect. *PLoS ONE*, *8*(7), e70026. https://doi.org/10.1371/journal.pone.0070026

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs. (Version 0.9.12-2). Retrieved from http://CRAN.R-project.org/package=BayesFactor

Murdock, B. B. (1965). Effects of a Subsidiary Task on Short-Term Memory. *British Journal of Psychology*, *56*(4), 413–419. https://doi.org/10.1111/j.2044-8295.1965.tb00983.x

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*, 53–81. https://doi.org/10.1146/annurev.psych.53.100901.135131

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

Oberauer, K. (2005). Binding and Inhibition in Working Memory: Individual and Age Differences in Short-Term Recognition. *Journal of Experimental Psychology: General*, *134*(3), 368–387. https://doi.org/10.1037/0096-3445.134.3.368

Oberauer, K. (2009). Design for a Working Memory. In *Psychology of Learning and Motivation* (Vol. 51, pp. 45–100). Elsevier. https://doi.org/10.1016/S0079-7421(09)51002-X

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*(3), 193–198. https://doi.org/10.1037/h0049234

Portrat, S., & Lemaire, B. (2015). Is Attentional Refreshing in Working Memory Sequential? A Computational Modeling Approach. *Cognitive Computation*, *7*(3), 333–345. https://doi.org/10.1007/s12559-014-9294-8

R core team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex*, *43*, 135–145. https://doi.org/10.1016/S0010-9452(08)70451-9

Ricker, T. J., & Cowan, N. (2010). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1355–1368. https://doi.org/10.1037/a0020356

Ricker, T. J., Vergauwe, E., Hinrichs, G. A., Blume, C. L., & Cowan, N. (2015). No Recovery of Memory when Cognitive Load is Decreased. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *41*(3), 872–880. https://doi.org/10.1037/xlm0000084

Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition*, *22*(5), 511–524. https://doi.org/10.3758/BF03198390

Rose, N. S., Buchsbaum, B. R., & Craik, F. I. M. (2014). Short-term retention of a single word relies on retrieval from long-term memory when both rehearsal and refreshing are disrupted. *Memory & Cognition*, *42*(5), 689–700. https://doi.org/10.3758/s13421-014-0398-x

Rose, N. S., & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal* *of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1019–1029. https://doi.org/10.1037/a0026976

Rose, N. S., Craik, F. I. M., & Buchsbaum, B. R. (2015). Levels of Processing in Working Memory: Differential Involvement of Frontotemporal Networks. *Journal of Cognitive Neuroscience*, *27*(3), 522–532. https://doi.org/10.1162/jocn_a_00738

Rose, N. S., Myerson, J., Roediger, H. L., & Hale, S. (2010). Similarities and differences between working memory and long-term memory: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 471–483. https://doi.org/10.1037/a0018405

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Souza, A. S., & Oberauer, K. (2017a). The contributions of visual and central attention to visual working memory. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-017-1357-y

Souza, A. S., & Oberauer, K. (2017b). Time to process information in working memory improves episodic memory. *Journal of Memory and Language*, *96*, 155–167. https://doi.org/10.1016/j.jml.2017.07.002

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: thinking of an item improves retrieval from visual working memory: Refreshing working memory traces. *Annals of the New York Academy of Sciences*, *1339*(1), 20–31. https://doi.org/10.1111/nyas.12603

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654. https://doi.org/10.1126/science.153.3736.652

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, *4*, 421–457.

Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(6), 1020–1034. https://doi.org/10.1037/0278-7393.33.6.1020

Unsworth, N. (2008). Exploring the Retrieval Dynamics of Delayed and Final Free Recall: Further Evidence for Temporal-Contextual Search. *Journal of Memory and Language*, *59*(2), 223–236. https://doi.org/10.1016/j.jml.2008.04.002

Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. *Acta Psychologica*, *134*(1), 16–28. https://doi.org/10.1016/j.actpsy.2009.11.010

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*(1), 68–80. https://doi.org/10.1016/j.jml.2005.06.003

Unsworth, N., & Engle, R. W. (2007a). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*(6), 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., & Engle, R. W. (2007b). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132. https://doi.org/10.1037/0033-295X.114.1.104

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do Mental Processes Share a Domain-General Resource? *Psychological Science*, *21*(3), 384–390. https://doi.org/10.1177/0956797610361340

Vergauwe, E., Camos, V., & Barrouillet, P. (2014). The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1072–1095. https://doi.org/10.1037/a0035779

Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval rate may describe many cognitive procedures. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00126

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychonomic Bulletin & Review*, *22*(4), 1001–1006. https://doi.org/10.3758/s13423-014-0755-6

Vergauwe, E., & Langerock, N. (2017). Attentional refreshing of information in working memory: Increased immediate accessibility of just-refreshed representations. *Journal of Memory and Language*, *96*, 23–35. https://doi.org/10.1016/j.jml.2017.05.001

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1207–1241. https://doi.org/10.1037/a0020122

White, K. G. (2012). Dissociation of short-term forgetting from the passage of time. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(1), 255–259. https://doi.org/10.1037/a0025197

Wickens, D., D., Moody, M., J., & Dow, R. (1981). The nature and timing of the retrieval process and of interference effects. *Journal of Experimental Psychology: General*, *110*(1), 1–20. https://doi.org/10.1037/0096-3445.110.1.1

Table 1. Means (and 95% within-subjects confidence intervals; Cousineau, 2005; Morey, 2008) for initial recall, both serial scoring and free scoring, across experiments as a function of list length and the number of distractors following the memoranda.

			Distractors				
Experiment	Measure	List Length	0	1	2	3	3 slow
1	Serial scoring	2	0.98 (0.05)	0.88 (0.07)	0.70 (0.07)	0.73 (0.07)	-
		3	0.96 (0.05)	0.66 (0.08)	0.58 (0.08)	0.53 (0.06)	-
		4	0.84 (0.07)	0.52 (0.06)	0.39 (0.08)	0.45 (0.07)	-
	Free Scoring	2	0.98 (0.04)	0.91 (0.05)	0.75 (0.06)	0.75 (0.06)	-
		3	0.96 (0.04)	0.78 (0.06)	0.73 (0.06)	0.66 (0.05)	-
		4	0.92 (0.04)	0.73 (0.04)	0.59 (0.06)	0.60 (0.05)	-
2	Serial Scoring	2	1.00 (0.06)	0.80 (0.05)	0.62 (0.06)	0.64 (0.08)	-
		4	0.77 (0.06)	0.46 (0.04)	0.38 (0.05)	0.31 (0.06)	-
	Free Scoring	2	1.00 (0.04)	0.86 (0.05)	0.73 (0.05)	0.72 (0.07)	-
		4	0.93 (0.03)	0.68 (0.03)	0.62 (0.04)	0.52 (0.06)	-
3	Serial Scoring	2	0.99 (0.04)	-	-	0.71 (0.06)	0.69 (0.04)
		4	0.83 (0.05)	-	-	0.33 (0.05)	0.32 (0.04)
	Free Scoring	2	0.99 (0.03)	-	-	0.77 (0.06)	0.76 (0.04)
		4	0.91 (0.03)	-	-	0.56 (0.04)	0.53 (0.04)

Figure 1. A schematic of the task used in the current experiments, the hypothetical psychological processes involved, and the expected pattern of results. WM = working memory, M = memoranda (i.e., in this study, the words to recall), D = distractors (i.e., in this study, the arithmetic problems following the words), FFR = final free recall.

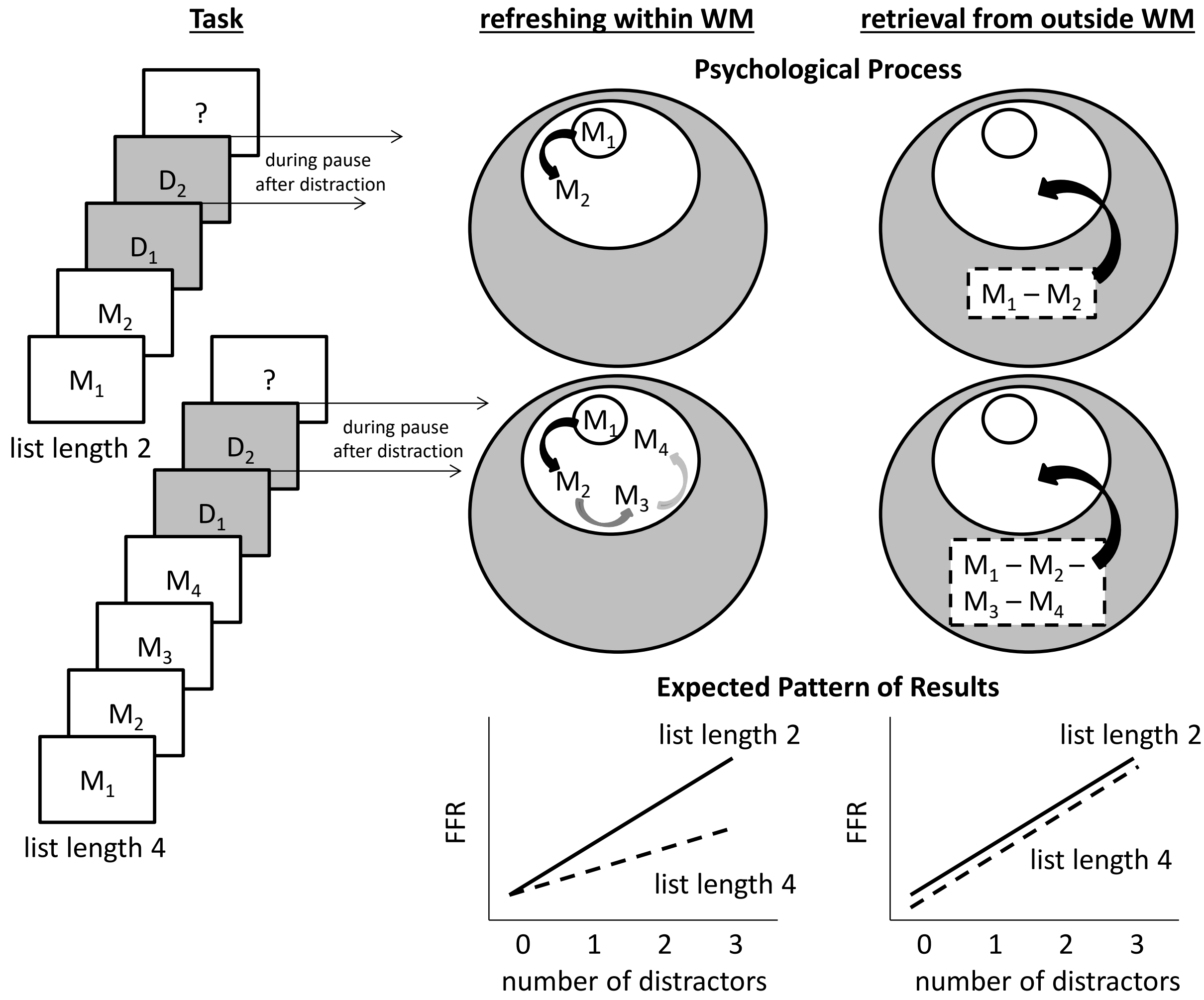

Figure 2. A. Initial recall – serial scoring, B. Initial recall – free scoring, C. Final free recall (FFR), and D. Final free recall conditionalized on initial recall (FFR*) across serial position in Experiment 1. Bars reflect 95% confidence intervals.

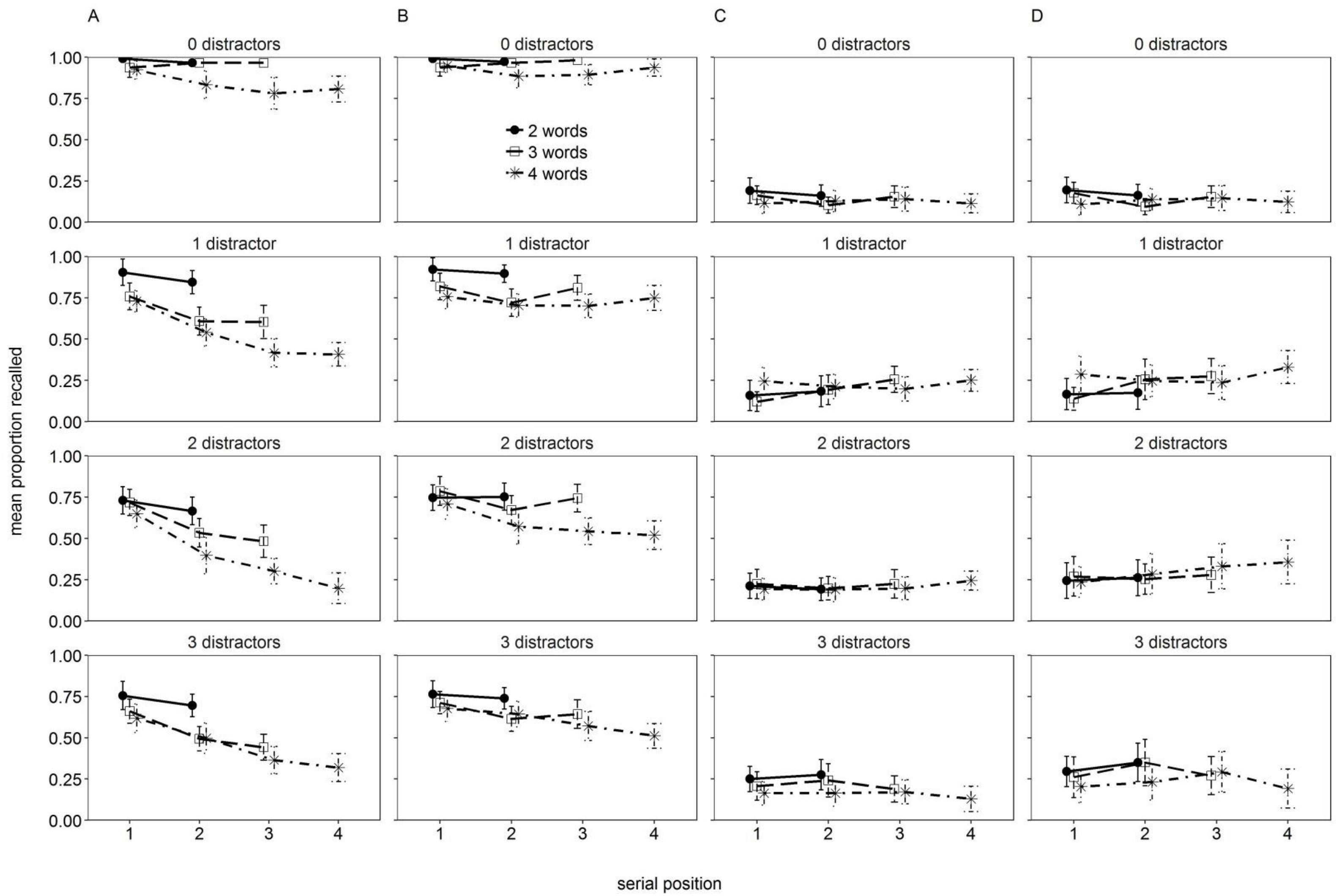

Figure 3. A. Final free recall (FFR) and B. FFR* (final recall conditionalized on initial recall) as a function of list length and distractors in Experiment 1. Bars reflect 95% within-subjects confidence intervals.

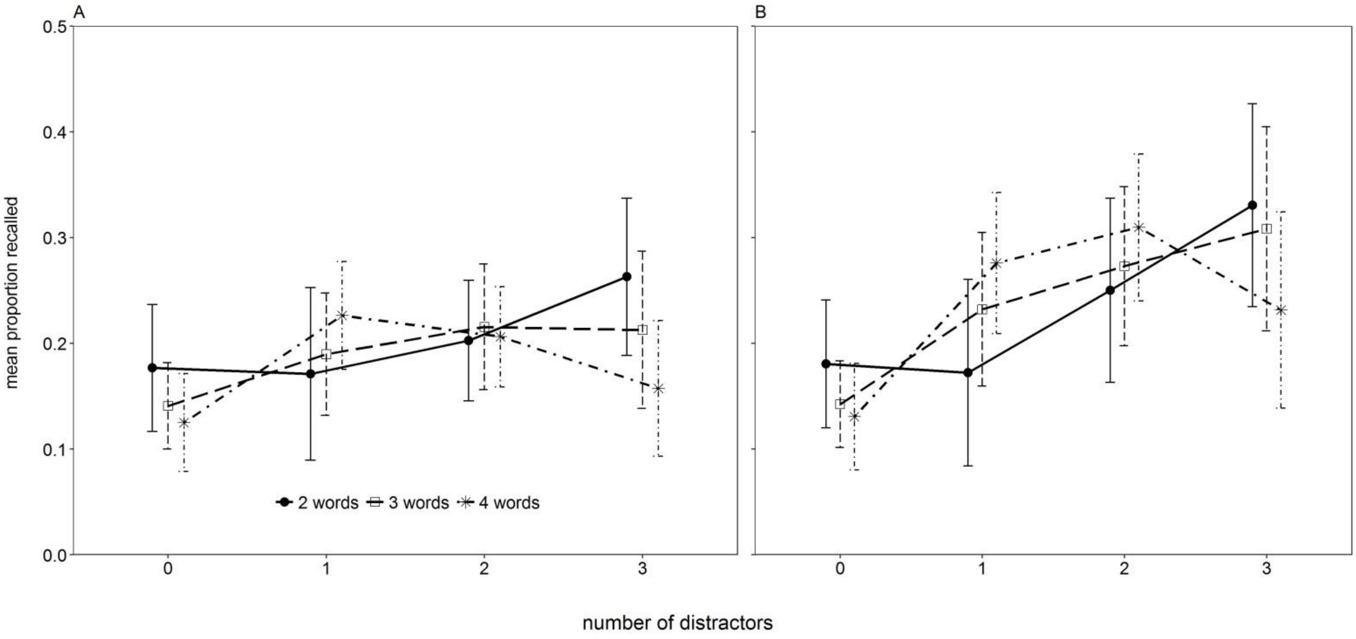

Figure 4. A. Initial recall – serial scoring, B. Initial recall – free scoring, C. Final free recall (FFR), and D. Final free recall conditionalized on initial recall (FFR*) across serial position in Experiment 2. Bars reflect 95% within-subjects confidence intervals.

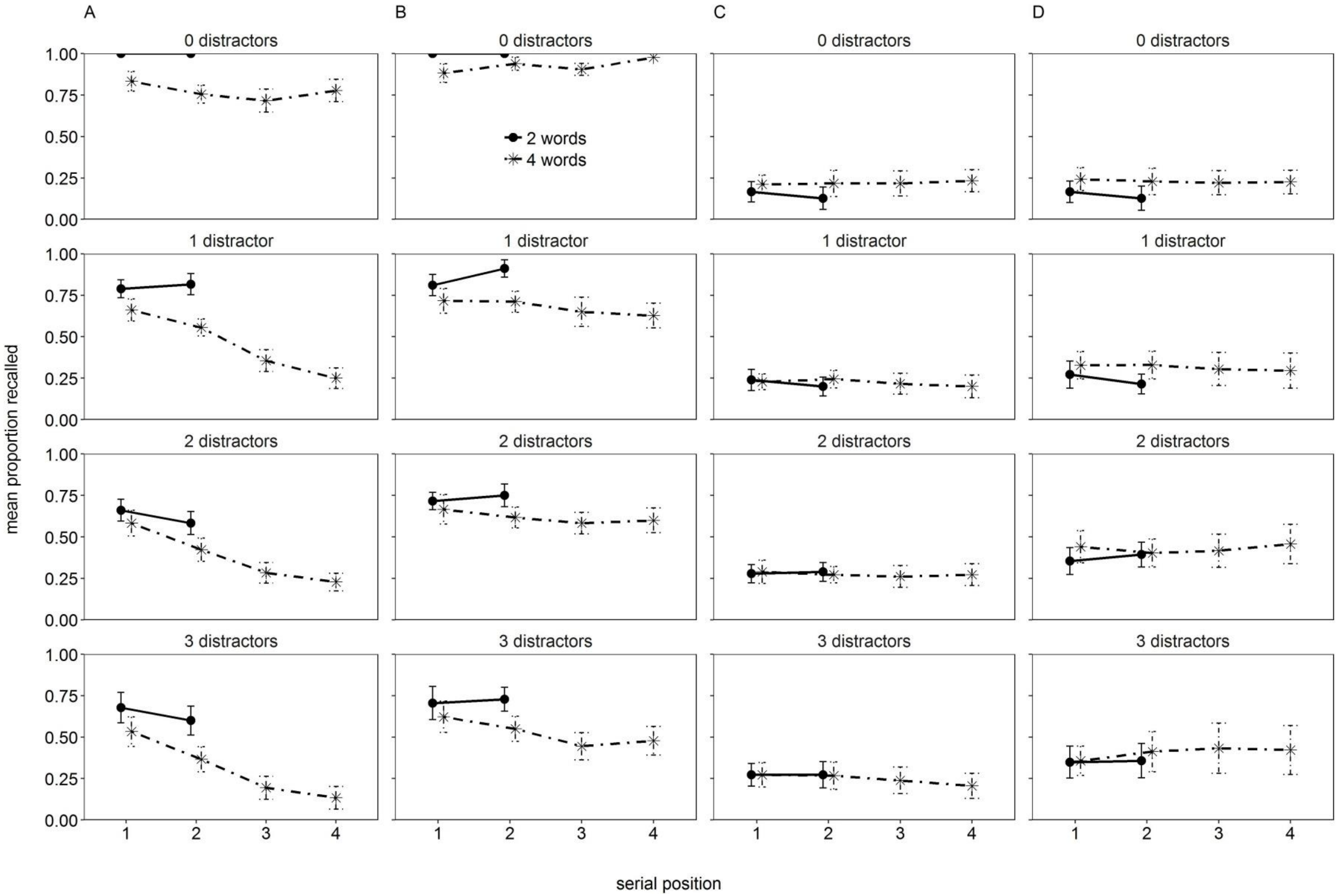

Figure 5. Final free recall (FFR) and B. FFR* (final recall conditionalized on initial recall) as a function of list length and distractors in Experiment 2. Bars reflect 95% within-subjects confidence intervals.

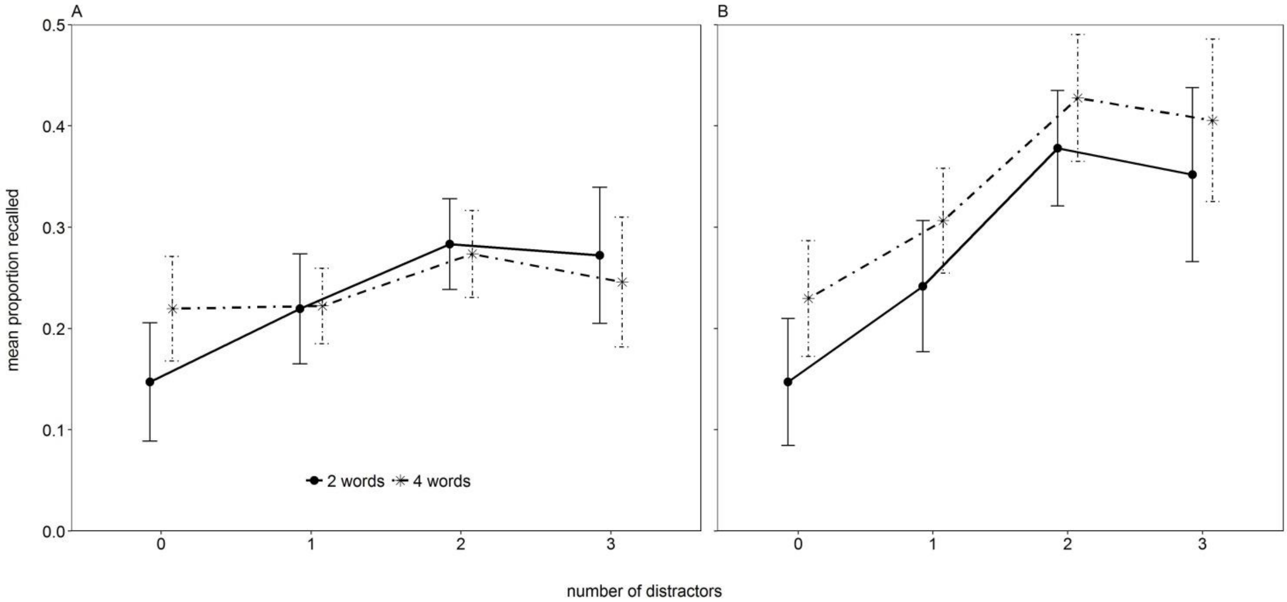

Figure 6. A. Initial recall – serial scoring, B. Initial recall – free scoring, C. Final free recall (FFR), and D. Final free recall conditionalized on initial recall (FFR*) across serial position in Experiment 3. Bars reflect 95% within-subjects confidence intervals.

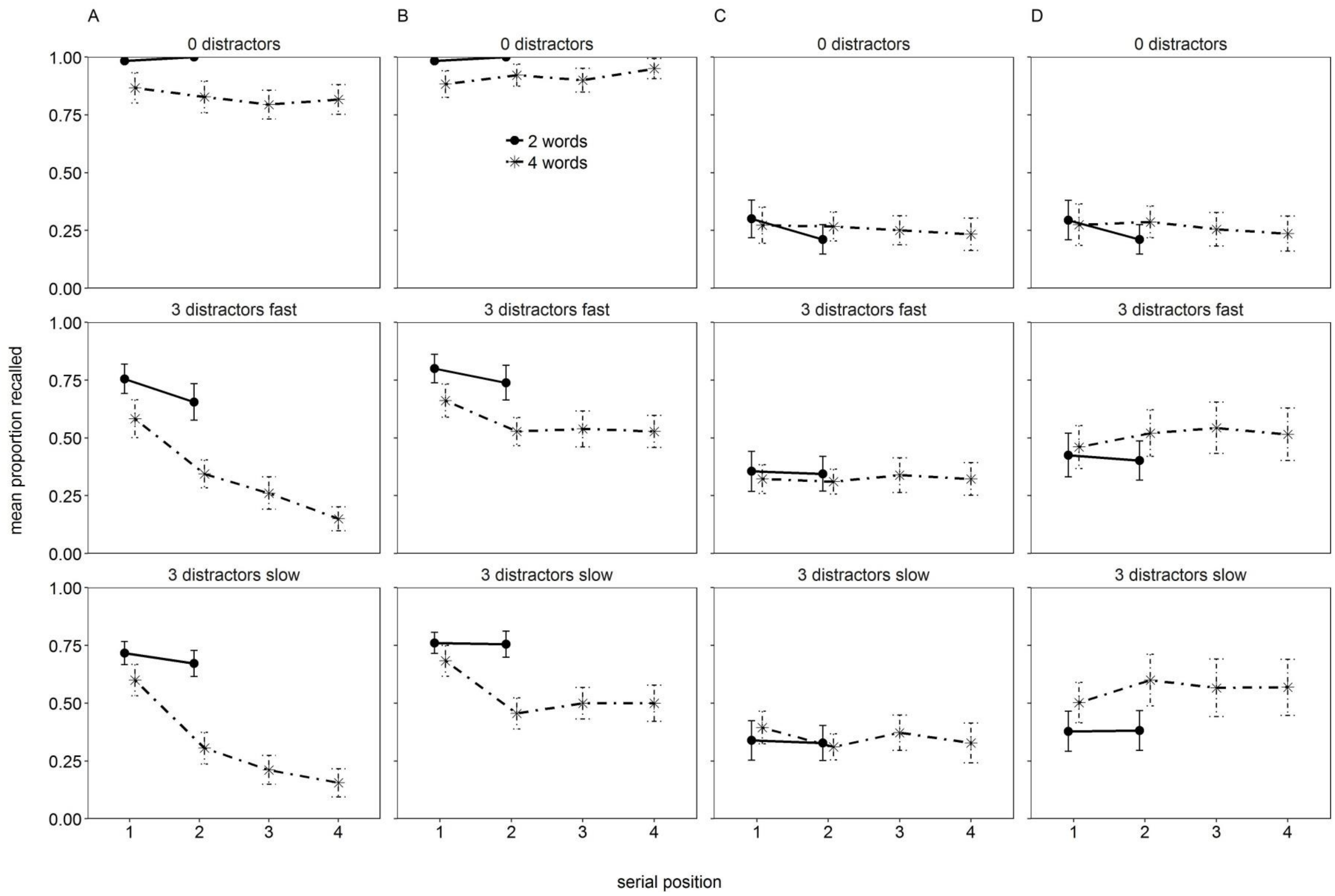

Figure 7. Final free recall (FFR) and B. FFR* (final recall conditionalized on initial recall) as a function of list length and distractors in Experiment 3. Bars reflect 95% within-subjects confidence intervals.