Journal of Memory and Language 104 (2019) 25–42

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

# Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation

Ashley L. Miller[a,*], Marina P. Gross[b], Nash Unsworth[a]

[a] Department of Psychology, University of Oregon, United States
[b] Department of Psychological & Brain Sciences, Washington University in St. Louis, United States

ARTICLE INFO

*Keywords:*
Attention
Working memory
Long-term memory
Individual differences
Recall

ABSTRACT

The present study used pupil dilation as an index of the intensity of attention to determine if variation in attention at encoding partially accounts for the relation between working memory capacity (WMC) and long-term memory (LTM). In Experiment 1, participants completed a delayed free recall task while pupil dilation was simultaneously recorded. Results revealed high WMC individuals displayed an increase in pupil dilation across serial positions, whereas low WMC individuals exhibited a decline in pupil dilation. Experiment 2 employed a similar method but manipulated encoding conditions via value–directed remembering. Results demonstrated when later serial positions were labeled as more important, the pupillary response no longer declined for low WMC individuals. Instead, low WMC individuals increased attention across serial positions, with the caveat being that these individuals devoted less attention than high WMC individuals to all items under these conditions. Overall, results support the notion that high WMC individuals outperform low WMC individuals in delayed free recall, which is partly explained by the amount of attention devoted to items at encoding.

## Introduction

Working memory (WM) is the ability to maintain and manipulate task relevant information in the presence of simultaneous processing and distraction (Baddeley & Hitch, 1974). WM is believed to encompass a resource limited system in which individuals can maintain approximately 4 ± 1 chunks of items (Cowan, 2001) in the current focus of attention and has been shown to predict a number of higher-order cognitive functions including reading and language comprehension (Daneman & Carpenter, 1980; Daneman & Merikle, 1996), general fluid intelligence (Ackerman, Beier, & Boyle, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen & Christal, 1990), and of particular interest to the current study, the ability to successfully encode and retrieve information from long term memory (LTM; Unsworth, 2010, 2016; Unsworth, Brewer, & Spillers, 2009). While higher-order cognitive functioning in general is important for various reasons, our ability to successfully remember information is essential to everyday functioning. Not only does LTM performance partially explain the relation between working memory capacity (WMC) and general fluid intelligence (Unsworth et al., 2009; Unsworth, 2009a, 2010), but on a daily basis we are faced with the task of remembering an impending deadline, previously learned facts necessary for an exam or one's job, the name of an acquaintance, and more. As such, encoding and retrieving relevant information is a critical component of navigating the world around us. Given the importance of encoding and retrieval of relevant information, it is imperative that researchers better understand why some people (e.g., high WMC individuals) are better at remembering information than others. The present study sought to further address this question.

### *WMC and LTM*

Research has demonstrated that high WMC and low WMC individuals differ in various aspects of LTM, including free (e.g., Unsworth, 2007) and cued (e.g., Unsworth, 2009b) recall. In prior work we have suggested a number of important reasons for these WMC related differences, including variation in overall search set size (i.e., search efficiency; Miller & Unsworth, in press; Unsworth, 2007; Unsworth & Engle, 2007) and variation in monitoring abilities (Unsworth & Brewer, 2010). Of note, these processes largely reflect control processes at retrieval. With respect to control processes at encoding, prior work suggests variation in encoding strategy use

---

* Corresponding author at: Department of Psychology, 1227 University of Oregon, Eugene, OR 97403, United States.
*E-mail address:* amiller8@uoregon.edu (A.L. Miller).

https://doi.org/10.1016/j.jml.2018.09.005
Received 4 January 2018; Received in revised form 18 September 2018
Available online 28 September 2018

(Unsworth, 2016) may also account for some of the WMC-LTM relationship. It remains to be seen, however, whether other factors that influence the strength of memory representations in LTM could likewise account for some of the results discussed previously. Recovery of items from LTM is determined by an item's absolute strength (Rohrer, 1996); hence recovery is likely if the strength of an item exceeds some critical threshold. One factor that may influence the strength of recoverable items is the amount of attention that item receives at encoding, such that items that receive more attention at encoding may have greater strengths (see Unsworth, 2009a for related discussion).

Research consistent with this view (e.g., Anderson, Craik, & Naveh-Benjamin, 1998; Baddeley, Lewis, Eldridge, & Thomson, 1984) shows that dividing attention at encoding significantly impairs recall performance on a variety of LTM tasks, including free recall and paired-associates tasks. That is, when attention is not fully devoted to encoding items, those items are weakly encoded and chances of recovery are low. The notion that attention is important for encoding has also been used to explain levels of processing effects (see Craik & Lockhart, 1972). In these cases, it is not necessarily the amount of attention (or time spent attending to stimuli) that determines subsequent episodic memory. Rather, it is the elaborative nature of attentional processing at encoding. In either case, lower probability of recall may be attributed to items having lower recoverable strengths, which may be due to those items receiving less attentional processing at encoding. If individual differences in WMC are related to how much attention individuals allocate to items at encoding, this may be another mechanism responsible for recall accuracy findings that researchers (e.g., Unsworth, 2016; Unsworth & Brewer, 2010) commonly associate with search efficiency and monitoring processes.

In support of this claim, Kane and Engle (2000) showed that dividing attention at encoding impaired recall performance more so for high WMC individuals than for low WMC individuals, suggesting high WMC individuals engage in more attentional processing under normal learning conditions. What is more, substantial evidence exists demonstrating the importance of attentional factors in accounting for individual differences in WMC, particularly in terms of attention control (Engle & Kane, 2004). Therefore, it seems possible that individual differences in WMC could be related to differences in how much attention is allocated to items at encoding. The present study sought to address this possibility and to see whether differences in this aspect of attention control could partly explain high WMC individuals' greater recall accuracy. As such, the particular mechanism of interest in the present study is the intensity of attention devoted to items at encoding, which may be indexed via pupillometry.

### *Pupillary response as an index of attention at encoding*

A great deal of prior research suggests that task evoked pupillary responses (TEPRs) reflect changes in pupil dilation relative to baseline levels due to the attentional demands imposed by a cognitive task (Beatty & Lucero-Wagoner, 2000; Goldinger & Papesh, 2012). For instance, the pupil dilates as math problem difficulty increases (Hess & Polt, 1964), as well as when memory load increases in traditional short term memory tasks (Kahneman & Beatty, 1966; Peavler, 1974). Research has further demonstrated that once memory load exceeds capacity limits, the pupillary response sometimes diminishes and displays a negative slope (Granholm, Asarnow, Sarkin, & Dykes, 1996; Granholm, Morris, Sarkin, Asarnow, & Jeste, 1997; Van Gerven, Paas, Van Merrienboer, & Schmidt, 2004), which is believed to occur once individuals are no longer able to or refuse to allocate additional resources to the task. More recent research (Unsworth & Robison, 2015) has also shown that individuals differentially allocate attention to items in WM as a function of the number of to-be-remembered items in a WM task. Specifically, during a delay period (after stimulus presentation and before recall) pupil dilation increased and reached an asymptote corresponding to the amount of items being maintained in one's WM. Results such as these led Kahneman (1973) to suggest that pupil dilation is a reliable and valid psychophysiological marker of attentional allocation. That is, TEPRs correspond to the intensive aspect of attention and provide an online indication of the amount of attentional effort devoted to a given item (i.e., the "intensity of attention"; Kahneman, 1973; Just & Carpenter, 1993).[1]

Using TEPRs, prior research has also linked pupillary responses at encoding to LTM performance (Ariel & Castel, 2014; Engle, 1975; Kafkas & Montaldi, 2011; Papesh, Goldinger, & Hout, 2012). For example, Ariel and Castel (2014) administered a value directed remembering task and found increased TEPRs for high value words relative to low value words. Notably, high value words were also associated with improved recall. Moreover, Papesh et al. (2012) demonstrated that the highest confident hits at test (i.e., items correctly recognized associated with the greatest confidence) were also associated with larger dilation during encoding. Thus, items that received the most attentional effort at encoding were more likely to be better remembered. While results such as these suggest the relation between TEPRs at encoding and ensuing LTM performance is *positive* in nature, it is important to acknowledge that the direction of this effect appears to be paradigm specific. Namely, using incidental learning conditions, Kafkas and Montaldi (2011) demonstrated the opposite pattern with pupil size when predicting recognition memory. Items that were subsequently remembered were associated with *decreased* TEPRs during encoding. Nonetheless, prior work adopting a similar procedure to ours (i.e., intentional learning conditions; Ariel & Castel, 2014; Goldinger, He, & Papesh, 2009; Papesh et al., 2012) collectively suggests that items associated with larger TEPRs at encoding receive more attentional effort, and these items are more likely to be better remembered.

The relation between pupil dilation at encoding and subsequent memory could be due, in part, to functioning of the locus coeruleus norepinephrine (LC-NE) neuromodulatory system, which is thought to be important for regulating attentional resources to maintain alertness and task engagement in a variety of situations (Aston-Jones & Cohen, 2005; Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Sara, 2009). Prior research has shown an important link between pupil dilation and the LC-NE (Gilzenrat et al., 2010; Murphy, Robertson, Balsters, & O'Connell, 2011; Sterpenich et al., 2006) and has suggested that pupil dilation during encoding provides an indirect index of LC-NE functioning (Eldar, Cohen, & Niv, 2013). The LC has direct projections to the hippocampus (Samuels & Szabadi, 2008), and it has been suggested that the LC is critically important for memory formation, potentially due to attentional modulation of hippocampal neurons (Rowland & Kentros, 2008). Thus, the LC-NE system may be particularly important for modulating the intensity of attention to items during encoding, which results in stronger hippocampal representations that are then easier to retrieve at recall. Critically, the functioning of the LC-NE system may also be a source of individual differences in WMC and attention control (Unsworth & Robison, 2017a). People with low WMC and/or low attention control abilities may suffer from a dysregulation of LC activity, such that these individuals exhibit more fluctuations in LC activity than high ability individuals. Given the role of the LC-NE system in both memory formation and attention control, it seems increasingly plausible that individual differences in WMC could relate to differences in how much attention is allocated to items at encoding.

---

[1] We do not mean to suggest that phasic pupil dilation always indexes the intensity of attention. Pupillary responses also reflect changes in luminance (i.e., pupillary light reflex; Binda, Pereverzeva, & Murray, 2013), arousal (e.g., Janisse, 1977; Phaf & Wolters, 1993), and more (e.g., Bijleveld, Custers, & Aarts, 2009; Braem, Coenen, Bombeke, van Bochove, & Notebaert, 2015). We attempted to control for these influences in our procedure outlined in the method section but caution the reader to note that other processes may also be at play.

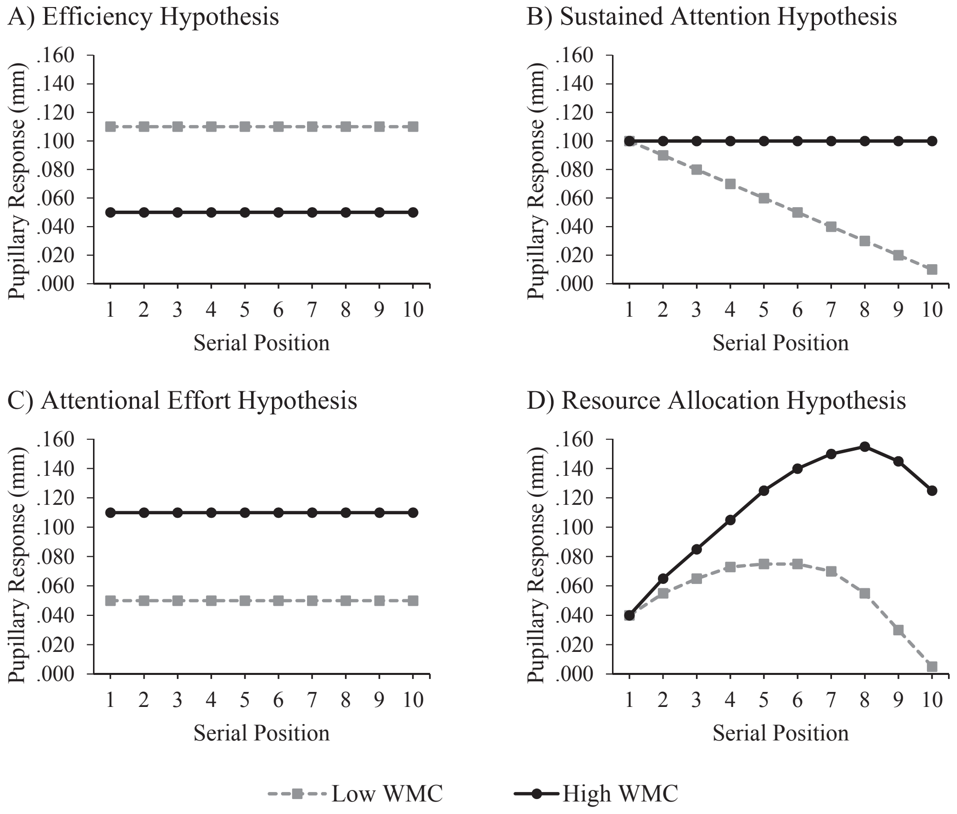

**Fig. 1.** Potential outcomes for how high and low WMC individuals may differ in intensity of attention throughout the encoding period across serial positions. Note that for the resource allocation hypothesis, memory capacity for high and low WMC individuals is assumed to be around 3–5 items (Cowan, 2001). But, because rehearsal and other strategies likely increase the number of items maintained during encoding, we estimated that high WMC individuals would be able to maintain 7 total items compared to 5 items for low WMC individuals. Points of overloading were calculated by obtaining the value equating to 125% of each of these limits (Granholm et al., 1996).

## Present study

Collectively, the above studies suggest that attentional influences play an important role in explaining not only individual differences in WMC, but LTM performance as well. While prior work further suggests attentional factors at encoding may account for the relation between WMC and LTM (Kane & Engle, 2000), it remains to be seen whether differences in WMC specifically relate to the intensity of attention at encoding. Consequently, the present study utilized pupil dilation as an online measure of the intensity of attention devoted to items at encoding to better address this question. Of course, it is possible that high and low WMC individuals do not differ in the amount of attention devoted to items at encoding, meaning we would expect no differences between high and low WMC individuals' TEPRs across all trials (*null hypothesis*). However, if a relation between WMC and the intensity of attention exists, we aimed to assess the validity of four potential reasons that could explain how differential attention allocation to items at encoding could arise.

As demonstrated in Fig. 1, one possibility is that high WMC and low WMC individuals may differ in the efficiency of processing at encoding. That is, if efficient encoding is characterized by attaining maximal recall performance with minimal effort during encoding and high WMC individuals encode items more efficiently, we would expect people with high WMC to recall more correct items while putting forth less attentional effort during encoding. On the other hand, if low WMC individuals encode items less efficiently than do high WMC individuals, we would expect these individuals to put forth extra work during encoding to obtain a level of recall performance that remains lower than the performance of people with high WMC. Thus, we would expect larger TEPRs for low WMC individuals relative to high WMC individuals across all trials (*efficiency hypothesis;* see Ahern & Beatty, 1979). A second possibility is that high WMC individuals' superior attention control abilities may facilitate their ability to more consistently sustain attention across the entire encoding period for a list, meaning we would expect consistent TEPRs across serial positions for high WMC individuals. However, since low ability individuals are less able to maintain high levels of attentional processing (Unsworth et al., 2009), we would expect a significant decline in TEPRs across serial positions for low WMC individuals (*sustained attention hypothesis*).

Alternatively, it is also conceivable that high WMC individuals may devote more attention at encoding in general when compared to people with low WMC; hence we would expect larger TEPRs for high WMC individuals across the entire encoding period of a list. Such a result may arise if high WMC individuals have more attentional resources available for encoding, or because high WMC individuals may be more motivated and therefore devote more attentional effort in general (*attentional effort hypothesis;* see Heitz, Schrock, Payne, & Engle, 2008; Van der Meer et al., 2010). Lastly, a related idea concerns differences in how individuals are allocating these attentional resources. As new to-be-remembered items are introduced at encoding, individuals may try to utilize effortful encoding strategies that allow them to maintain all previously seen words in addition to the newly introduced word. As each new word is incorporated into the ongoing strategy, attentional effort and memory load would be expected to increase (see Kahneman & Beatty, 1966). Because high WMC individuals have greater maximum capacities that allow them to maintain more items in the current focus of attention (Cowan et al., 2005; Unsworth, Spillers, & Brewer, 2010), they should demonstrate an overloading function later than what would be observed for low WMC individuals. Therefore, we would expect both WMC groups to display increases in TEPRs until capacity limits are reached, with TEPR differences between high and low WMC individuals growing as task difficulty increases (i.e., as memory load increases; see Van der Meer et al., 2010). At or near capacity limitations, an asymptote in TEPRs would then occur until processing load exceeds 125% of maximum capacity (Granholm et al., 1996). Once processing load reaches these points of overloading, a decline in pupil dilation would be expected to follow (*resource allocation hypothesis*; Granholm et al., 1996; Van der Meer et al., 2010).

To better illuminate these possibilities, Experiment 1 required participants to complete three measures of WMC followed by a delayed free recall task, during which pupil dilation was simultaneously recorded. Experiment 2 attempted to test a hypothesis derived from the results of Experiment 1 by similarly using a delayed free recall task while pupil dilation was recorded. The primary difference between methodologies was that Experiment 2 implemented a value-directed remembering manipulation at encoding in an attempt to force participants to attend more to items presented either at the beginning or end of a list. Assessing individual differences in WMC while further examining how various encoding conditions influence the intensity of attention during encoding will allow us to better clarify WMC related differences in LTM abilities.

## Experiment 1

Experiment 1 had four primary goals. First, we sought to replicate the finding that pupillary responses at encoding reflect the amount of attention devoted to each item, which has been shown to relate to memory performance (e.g., Ariel & Castel, 2014; Papesh et al., 2012). Experiment 1 aimed to expand this research by secondly examining whether differences in the intensity of attention, as indexed by TEPRs, are related to individual differences in WMC. If WMC related differences do exist in the amount of attention devoted to items at encoding, we sought to differentiate which of the four previously mentioned hypotheses best accounts for the results. Third, another aim of the present study was to determine whether encoding strategy use is an additional mechanism by which WMC related differences in TEPRs arise. As previously mentioned, prior research has demonstrated that WMC related differences exist in encoding strategy use (Dunlosky & Kane, 2007; McNamara & Scott, 2001; Turley-Ames & Whitfield, 2003) and that encoding strategies partly explain the relation between WMC and episodic memory (e.g., Bailey, Dunlosky, & Kane, 2008). Of interest to the present study, the reason for this relation may be that effective encoding strategies are more resource demanding than ineffective encoding strategies. More specifically, a potential reason why ineffective strategies, such as the rehearsal strategy, are ideal for low WMC individuals is that these strategies could require the implementation of less attentional resources when compared to effective strategies (see Turley-Ames & Whitfield, 2003). If the pupil dilates in response to increased attentional demands, a strategy that requires increased attentional resources for successful implementation may be reflected in the pupillary response. Therefore, one may expect reports of effective encoding strategy use to be positively correlated with TEPRs, WMC, and recall accuracy.

As such, a final question we sought to address was how encoding strategy use, the intensity of attention, and WMC together predict recall performance. If a relationship exists between the intensity of attention and WMC, a primary question of interest was whether the intensity of attention partially accounts for the relation between individual differences in WMC and subsequent episodic memory performance. Consequently, a critical point of analysis was to also determine whether WMC and TEPRs at encoding share common variance in predicting recall accuracy. Moreover, if TEPRs explain some of the relation between WMC and recall accuracy and significant positive relations exist between WMC, TEPRs, and effective encoding strategy use, it is possible that effective encoding strategy use may be an additional factor that can explain how high WMC and low WMC individuals differentially allocate the intensity of attention to items at encoding. To further elucidate these possibilities, participants completed a delayed free recall task while pupil diameter was simultaneously recorded followed by a self-report encoding strategy questionnaire.

## Method

### *Participants*

Participants ($N = 139$) were recruited from the University of Oregon's human subject pool. Participants were between the ages of 18 and 29 and were compensated with course credit necessary for meeting a course research requirement. Data collection took place over two academic quarters. One participant had to be excluded from all analyses due to excessive missing pupil data on the delayed free recall task, leaving a total sample of 138 participants.

### *Materials and procedure*

After obtaining informed consent and demographic information, all participants were tested individually and completed three shortened measures of WMC: the operation span task (Ospan), the symmetry span task (Symspan), and the reading span task (Rspan). After completion of the WMC tasks, participants were moved to a dimly lit room where they completed a delayed free recall task while pupil diameter was simultaneously recorded binocularly at 120 Hz using a Tobii T120 eye-tracker. Prior to beginning the recall task, participants were seated roughly 60 cm from the monitor, and a 6-point standard calibration procedure began. To calibrate the eye tracker, participants were asked to fixate on a series of 6 grey dots presented on a white background. The Tobii Eye Tracker measures aspects of the participant's eyes and uses them together with an internal, anatomical 3D eye model to calculate the mapping between the identified gaze position on the display and the eye tracker's estimate of that position. Recalibration occurred whenever the criterion defined by the proprietary software was not met. All participants were successfully calibrated within the first few attempts. Upon completion of the delayed free recall task, participants were presented with a strategy report questionnaire on computer. Note that participants completed the tasks reported here as part of a larger experimental test battery. Since the other tasks administered during the experimental session do not relate to the current study, they are not reported. The entire experimental session lasted approximately one and a half hours.

### *WMC tasks*

**Ospan.** Participants solved a series of elementary math problems while remembering unrelated letters. First, on computer participants were presented with a math operation (e.g., $(4 \times 1) + 2 = ?$) in which they had to click the mouse to indicate that they had solved the problem. A new screen then appeared with an answer to the math solution (e.g., 6), whereby participants had to indicate if the answer listed onscreen was correct of incorrect via mouse click (e.g., in the case above, the answer 6 would be correct). Upon completion of the math operation, participants were then presented with a letter (e.g., F, H, J, K, L, N, P, Q, R, S, T, and Y) for 1 s. Immediately following letter presentation, the next math problem was presented. Set sizes varied randomly from 3 to 7 math operation/letter strings, and participants had to complete 2 trials of each set size for a total possible score of 50. At recall for each set, letters from the corresponding set had to be recalled in order by selecting the relevant letters. See Unsworth, Heitz, Schrock, and Engle (2005) for more details.

**Symspan**. Participants solved symmetry judgments while remembering the location of a sequence of red squares within a matrix. Symmetry judgments consisted of an $8 \times 8$ matrix of squares in which some of the squares were filled black and the remaining squares remained white. Participants indicated whether the pattern created by the filled squares was symmetrical about the vertical axis. Once participants indicated whether they believed the pattern was symmetrical or non-symmetrical, participants were shown a $4 \times 4$ matrix with one of the cells filled red for 650 ms. Immediately following the presentation of the red square matrix, the next symmetry judgment trial began. Set sizes randomly ranged from 2 to 5, and there were 2 trials of each set size for a total possible score of 28. Participants were asked to recall the sequence of red-square locations based on the order in which they were presented across the corresponding trial. Participants indicated the appropriate location of each red-square by clicking on cells in an empty

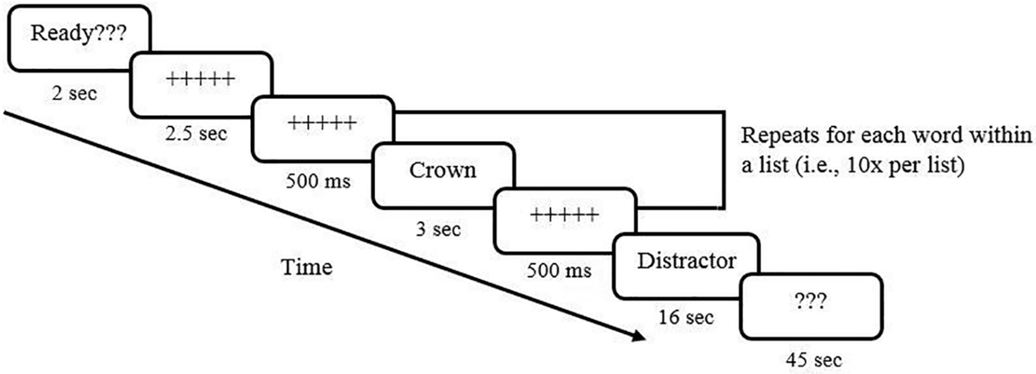

**Fig. 2.** Schematic of the experimental task.

matrix. See Unsworth, Redick, Heitz, Broadway, and Engle (2009) for more details.

**Rspan**. While remembering the same unrelated letters as in the Ospan, participants provided judgments about a series of sentences. More specifically, participants read a sentence containing 10–15 words and determined whether or not the sentence made sense to them (e.g., “Every now and then I catch myself swimming blankly at the wall”). Nonsense sentences were created by modifying a single word from an otherwise ordinary sentence (e.g., changing “staring” to “swimming” in the case above). Upon indicating whether the sentence made sense or not, participants were then presented with a letter for 1 s. Set sizes randomly varied from 3 to 7 sentence/letter strings, and participants had to complete 2 trials of each set size for a total possible score of 50. At recall for each set, letters from the corresponding set had to be recalled in order by selecting the appropriate letters. See Unsworth et al. (2009) for more details.

### *Delayed free recall task*

After calibration of the eye-tracker, participants were administered a delayed free recall task consisting of 5 word lists containing 10 words each. Word lists were initially composed of randomized nouns selected from the Toronto word pool (Friendly, Franklin, Hoffman, & Rubin, 1982), and all words were between 3 and 5 letters in length. Words (as well as the mask preceding/following each word) were presented in black text in Arial font (font size = 24) on a light grey background. Properties such as ambient light, screen brightness, contrast, etc. were held constant across participants. All participants received the same lists of words and were instructed to recall as many words as possible from each list. The task began with a “Ready?” signal onscreen for 2 s, followed by a fixation period lasting 2.5 s (baseline pupil diameter). Each list began with the same “Ready?” signal and fixation period, which were followed by a series of words presented individually in the center of the screen for 3 s. Each word was preceded and followed by a mask of 5 plus signs (e.g., “+ + + + +”) for 500 ms. After list presentation, participants then completed a 16 s distractor task that required participants to verbally report a series of 8 three-digit numbers in descending order (adapted from Rohrer & Wixted, 1994). Each 3-digit string was presented onscreen for 2 s. At recall, 3 question marks appeared in the center of the screen to prompt participants to recall as many words as possible within a 45 s window. Participants typed their responses in any order they wished and pressed “enter” after each word, thereby clearing the screen. See Fig. 2 for a schematic outlining the sequence of the experimental task.

### *Strategy report questionnaire*

Upon completion of the last delayed free recall list, participants reported on computer whether or not they used any encoding strategies to help better remember the words. Specifically, participants were shown the following options: (1) Read each word as it appeared, (2) Repeated the words as much as possible, (3) Used a sentence to link the words together, (4) Developed mental images of the words, (5) Grouped the words in a meaningful way, and (6) Did something else. Participants responded by typing their answers and were allowed to select more than one strategy. This strategy report questionnaire was based on similar reports used by Bailey et al. (2008) and Unsworth (2016).[2] Thus, ineffective strategies were characterized as passive reading and rehearsal, whereas effective strategies were characterized as interactive imagery, sentence generation, and grouping. Consistent with this work (Bailey et al., 2008; Bailey, Dunlosky, & Hertzog, 2009; Dunlosky & Kane, 2007; Unsworth, 2016), the “other” strategy category was excluded from our analyses because no a priori hypotheses were made as to whether these strategies would be more or less effective. While strategy reports of this type are less common, it seems feasible that this category may encompass mnemonic devices, such as the peg word system, method of loci, or rhyming (see Turley-Ames & Whitfield, 2003).

## Results

First, we created a single WMC composite score for each participant by averaging participants' standardized *Z*-scores on the Ospan, Symspan, and Rspan. This WMC composite score was used in all following analyses concerning WMC. Note that WMC was treated as a continuous variable in all analyses, meaning WMC was utilized as a categorical variable for graphical purposes only. Using a quartile split, the uppermost 25% of performers on the WMC tasks were categorized as high WMC, whereas the lowermost 25% performers were categorized as low WMC. Descriptive statistics are listed in Table 1. All measures demonstrated adequate variability around the mean, and assessments of skew and kurtosis were within normal ranges. Moreover, reliability estimates were satisfactory. Shown in Table 2 are the proportions of reported strategy use for each possible strategy. Correlations between all measures are displayed in Table 3.

### *Behavioral effects*

*Recall accuracy.* We first submitted recall accuracy to a repeated measures analysis of variance (ANOVA) with serial position (serial

---

[2] We probed for strategy use following task completion to avoid potential reactivity effects associated with concurrent strategy reports (see Dunlosky & Hertzog, 2001). While these retrospective reports correlated with overall recall performance (see Table 2), these reports are ill-suited in detecting whether specific strategies used on a given item or list relate to TEPRs and recall for said item. Such inferences would require either retrospective itemized strategy reports or concurrent strategy reports.

**Table 1**
Descriptive statistics and reliability estimates for all measures.

| Measure | *N* | *M* | *SD* | Skew | Kurtosis | Reliability |
|---|---|---|---|---|---|---|
| Ospan | 136 | 38.80 | 7.39 | −1.21 | 3.06 | .59 |
| Symspan | 138 | 19.35 | 4.95 | −.49 | −.36 | .54 |
| Rspan | 138 | 38.06 | 7.61 | −.41 | .02 | .64 |
| Accuracy | 138 | .57 | .16 | .32 | .13 | .85 |
| Mean TEPR | 138 | −.02 | .14 | −.12 | 2.48 | .91 |
| Ineffective | 137 | .71 | .39 | −.89 | −.79 | |
| Effective | 137 | .51 | .32 | −.12 | −.90 | |

*Note.* Two people were missing Ospan data because of a computer program malfunction, and one other person was missing strategy report data again due to a computer program malfunction.

**Table 2**
Proportions of reported strategy use as a function of strategy.

| Read | Repetition | Imagery | Sentence | Grouping | Other |
|---|---|---|---|---|---|
| .76 (.04) | .65 (.04) | .51 (.04) | .44 (.04) | .57 (.04) | .17 (.03) |

*Note.* One person excluded due to missing strategy data ($N = 137$). Proportions of strategies sum to greater than 1.0 because participants were allowed to report more than one strategy.

**Table 3**
Correlations among all measures.

| Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. WMC | – | | | | | | | |
| 2. Ospan | .80*** | – | | | | | | |
| 3. Symspan | .70*** | .34*** | – | | | | | |
| 4. Rspan | .77*** | .48*** | .25** | – | | | | |
| 5. Accuracy | .22** | .10 | .17* | .22* | – | | | |
| 6. Mean TEPR | .18* | .19* | .11 | .10 | .18* | – | | |
| 7. Ineffective | .06 | .10 | .03 | .02 | −.30*** | .06 | – | |
| 8. Effective | .01 | −.01 | .03 | −.01 | .45*** | .02 | −.14 | – |

* $p < .05$.
** $p < .01$.
*** $p < .001$.

positions 1 through 10) as a within-subjects factor. The analysis revealed a significant main effect of serial position, $F(9, 1233) = 96.51$, $MSE = .04$, $p < .001$, partial $\eta^2 = .41$, indicating words presented at earlier serial positions were more likely to be correctly recalled than words presented at later serial positions. This notion is further supported by polynomial contrasts, which revealed significant negative linear ($F(1, 137) = 381.57$, $MSE = .08$, $p < .001$, partial $\eta^2 = .74$) and quadratic ($F(1, 137) = 118.53$, $MSE = .05$, $p < .001$, partial $\eta^2 = .46$) trends, reflecting a strong primacy effect and a weak recency effect.

Next, we added WMC as a covariate to the above model. The repeated measures ANCOVA yielded a significant effect of WMC, $F(1, 136) = 6.92$, $MSE = .25$, $p = .009$, partial $\eta^2 = .05$, suggesting that higher WMC was related to better recall performance ($r = .22$, $p < .01$). Moreover, the interaction between WMC and serial position was likewise significant, $F(9, 1224) = 1.98$, $MSE = .04$, $p = .038$, partial $\eta^2 = .014$, meaning the effect of serial position on recall accuracy was influenced by WMC differences. Fig. 3 suggests both WMC groups tended to recall primacy (words at serial positions 1–3) items the most. However, low WMC individuals appeared to show no significant difference in proportion correctly recalled for mid (words at serial positions 4–7) and recency (words at serial positions 8–10) items, whereas high WMC individuals appeared to recall more mid items than the majority of recency items (with exception of a boost in recall for words at serial position 9).

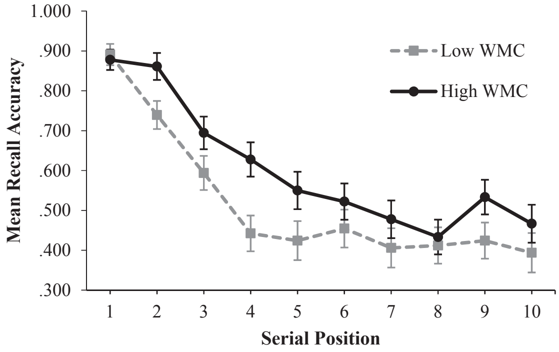

**Fig. 3.** Recall accuracy as a function of serial position for high WMC ($n = 36$) and low WMC ($n = 33$) individuals.

### *Pupillary effects*

Next, we examined pupillary responses, the primary analyses of interest. Pupil diameter was assessed continuously throughout the delayed free recall task. Data from each participant's left eye was used for analyses, and missing data points associated with eye tracker malfunction, blinks, or off-screen fixations were excluded from averaging (i.e., we did not interpolate missing pupil data). TEPRs were baseline corrected by subtracting mean baseline diameter from the average pupil diameter during the 3 s encoding phase for each word (word presentation).[3] In addition, the pupil data for the 3 s encoding phase for each word was broken down into a series of 200 ms timeframes, resulting in 15 total baseline corrected bins.

*Serial position and time course.* First, mean TEPRs during the encoding period were assessed as a function of time across the encoding period for each word and serial position. Note that an additional 9 participants were excluded from the serial position and time course pupil analyses, leaving a total sample of 129 participants. These individuals were excluded due to missing data across various serial positions. The 10 (serial position; within-subjects factor) × 15 (200 ms bin; within-subjects factor) repeated measures ANOVA revealed a significant main effect of bin, $F(14, 1792) = 23.79$, $MSE = .02$, $p < .001$, partial $\eta^2 = .16$. Fig. 4 shows that pupil diameter increased throughout the encoding period for each word. While part of the dilation across bins (particularly the first few bins) is likely just visual processing to the words, we don't believe these TEPRs are solely an artifact of visual processing. There was an additional significant main effect of serial position, $F(9, 1152) = 8.30$, $MSE = .13$, $p < .001$, partial $\eta^2 = .06$, indicating pupil diameter was smallest for primacy (items at serial positions 1–3) and recency (items at serial positions 8–10) items but largest for mid list items (items at serial positions 4–7), quadratic trend: $F(1, 128) = 33.22$, $MSE = .22$, $p < .001$, $\eta^2 = .21$. A significant two-way interaction between serial position and bin also emerged, $F(126, 16128) = 2.37$, $MSE = .01$, $p < .001$, partial $\eta^2 = .02$, suggesting pupil diameter was largest across the encoding period for mid list items and moderate for primacy and recency items.

To assess any influence of WMC on the aforementioned effects, we next added WMC as a covariate. The 10 (serial position) × 15 (200 ms bin) repeated measures ANCOVA revealed a significant effect of WMC, $F(1, 127) = 5.27$, $MSE = 3.20$, $p = .023$, partial $\eta^2 = .04$, suggesting

---

[3] In both experiments, WMC was unrelated to baseline pupil diameter (Experiment 1 $r = .09$, $p = .30$; Experiment 2 $r = -.13$, $p = .14$). While some prior work (Heitz et al., 2008; Tsukahara, Harrison, & Engle, 2016) has demonstrated a positive relation between WMC and baseline pupil, the observed null result is largely consistent with prior work in our lab (see Unsworth & Robison, 2015, 2017b). Note that the results reported herein focus on TEPRs, which we baseline corrected on a list-by-list basis. So the results do take into account potential differences in baselines.

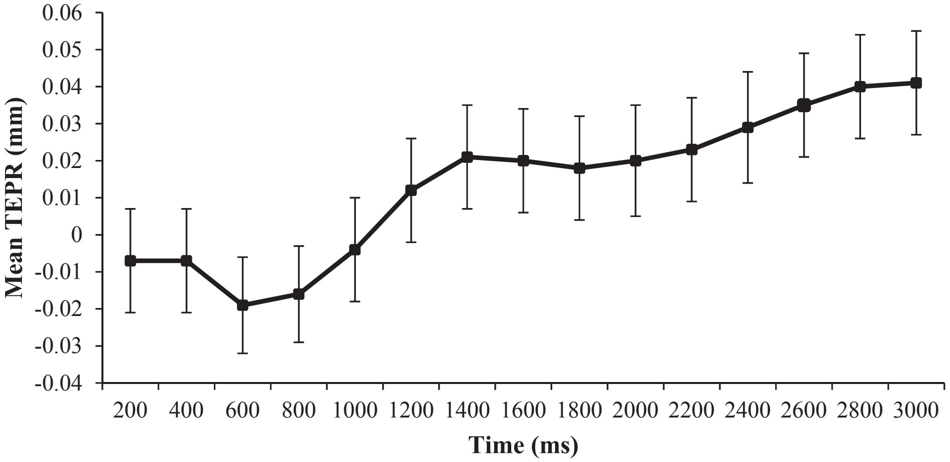

**Fig. 4.** Pupil diameter as a function of time point (bin) at encoding of each word.

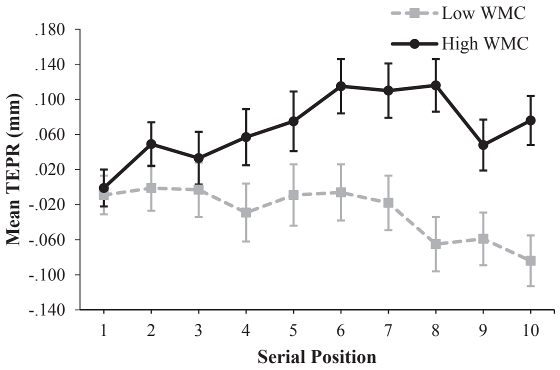

**Fig. 5.** Pupil diameter as a function of serial position for high WMC ($n$ = 33) and low WMC ($n$ = 31) individuals.

higher WMC was related to larger TEPRs ($r = .18$, $p = .037$). Importantly, there was also a significant serial position × WMC interaction, $F(9, 1143) = 4.82$, $MSE = .12$, $p < .001$, partial $\eta^2 = .04$, meaning the effect of serial position on pupil diameter differed as a function of one's WMC. Fig. 5 reveals that high WMC and low WMC individuals displayed no significant differences in pupil diameter across primacy items. After serial position 3, though, high WMC individuals showed an increase in pupil dilation across the remaining serial positions [positive linear trend: $F(1, 32) = 8.23$, $MSE = .31$, $p = .007$, partial $\eta^2 = .21$; quadratic trend: $F(1, 32) = 8.02$, $MSE = .27$, $p = .008$, partial $\eta^2 = .20$], whereas low WMC individuals showed a decrease in pupil dilation, negative linear trend: $F(1, 30) = 13.31$, $MSE = .20$, $p = .001$, partial $\eta^2 = .31$. As such, the correlation between WMC and TEPR was weakest for primacy items (items at serial positions 1–3; $r = .05$, $p > .05$), moderate for mid items (items at serial positions 4–7; $r = .18$, $p < .05$), and strongest for recency items (items at serial positions 8–10; $r = .25$, $p < .01$).

The ANCOVA further revealed a significant bin × WMC interaction, $F(14, 1778) = 1.86$, $MSE = .02$, $p < .05$, partial $\eta^2 = .014$, indicating high WMC individuals displayed larger pupil dilation throughout the encoding period for each word. There was also a significant three-way interaction between serial position, bin, and WMC, $F(126, 16002) = 1.50$, $MSE = .01$, $p < .001$, partial $\eta^2 = .012$. Fig. 6 reveals that for high WMC individuals, pupil dilation continued to gradually increase throughout the encoding period for all serial positions, with primacy items displaying smaller dilations than mid and recency items. Conversely, low WMC individuals showed moderate increases in dilation that appeared to plateau near the middle of the encoding period. Moreover, pupil dilation appeared to be largest for primacy items and smallest for recency items, despite a gradual increase in dilation for recency items.

*Regressions*. Lastly, we examined whether the pupillary response at encoding in conjunction with WMC and encoding strategy use predicted recall performance. We submitted WMC, mean TEPR, ineffective strategy use, and effective strategy use to a simultaneous linear regression model predicting recall accuracy. As seen in Table 4, all of the predictors together accounted for 33% of the variability in recall accuracy, $F(4, 132) = 16.25$, $p < .001$, with each of the predictors accounting for unique variance. WMC, mean TEPR, and effective strategy use were positively related to recall performance, meaning higher scores on these measures were associated with improved recall accuracy. On the other hand, ineffective strategy use was negatively related to recall performance, so use of more ineffective strategies was related to decreased recall accuracy. Altogether, these results suggest that individual differences in recall accuracy are partially driven by individual differences in WMC, the intensity of attention devoted to items at encoding, and use of ineffective and effective encoding strategies.[4]

To further examine the shared and unique contribution of WMC and mean TEPR with recall accuracy, we utilized variance partitioning methods (e.g., Chuah & Maybery, 1999; Cowan et al., 2005; Unsworth, Fukuda, Awh, & Vogel, 2014) to distribute the overall $R^2$ of recall accuracy into portions shared and unique to WMC and mean TEPR. Note, because a primary question of interest was to determine whether the intensity of attention (as indexed by mean TEPR) influences the WMC—LTM relationship, we only included these variables in the variance partitioning analyses. In addition, neither ineffective nor effective strategy use correlated with WMC or mean TEPR (all $p$'s $> .45$), so there was no reason to further explore shared variance with these factors. A series of regression analyses were used to obtain $R^2$ values from each of the predictors in order to partition the variance. For each variable entering the regression, the zero-order correlations from Table 3 were used.

As demonstrated in Fig. 7, the results suggest that a total of 6.8% of

---

[4] We also examined subsequent memory effects to see if pupillary responses during encoding would predict subsequent recall. We averaged TEPRs during encoding for recalled and forgotten items separately. Results revealed that TEPRs for subsequently recalled items in Experiment 1 were not significantly different from TEPRs for subsequently forgotten items, $t(135) = 1.54$, $p = .126$. Experiment 2 revealed an overall non-significant subsequent memory effect, $F(126) = 2.17$, $p = .143$, partial $\eta^2 = .017$, but this was qualified by an interaction with list type. TEPRs for subsequently recalled items ($M = .054$, $SE = .016$) were significantly larger than TEPRs for subsequently forgotten items ($M = .029$, $SE = .017$) on ascending lists only, $F(126) = 3.93$, $p = .05$, partial $\eta^2 = .03$.

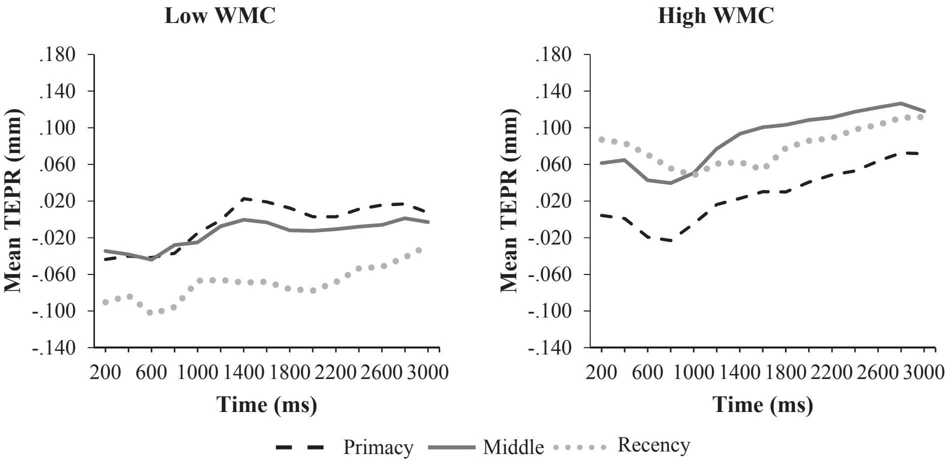

**Fig. 6.** Pupil diameter as a function of serial position and bin (time across encoding period) for low WMC ($n$ = 31) and high WMC ($n$ = 33) individuals. Serial position was broken down into Primacy (items 1–3), Mid (items 4–7), and Recency (items 8–10) for graphical purposes only.

**Table 4**
Simultaneous regression predicting recall accuracy.

| Variable | β | *t* | *sr*$^2$ | *R*$^2$ | *F* |
|---|---|---|---|---|---|
| WMC | .20 | 2.77$^{**}$ | .04 | | |
| Mean TEPR | .15 | 2.02$^{*}$ | .02 | | |
| Ineffective | −.26 | −3.66$^{***}$ | .07 | | |
| Effective | .41 | 5.62$^{***}$ | .16 | .33 | 16.25$^{***}$ |

*Note.* One person excluded from analysis because they were missing strategy report data ($N$ = 137).

* $p < .05$.

** $p < .01$.

*** $p < .001$.

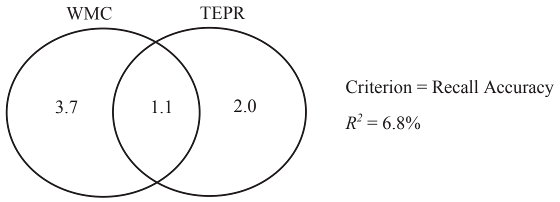

**Fig. 7.** Venn diagrams representing the shared and unique variance between WMC and mean TEPR in predicting recall accuracy.

the variance in recall accuracy were accounted for by WMC and the intensity of attention to items at encoding. WMC uniquely accounted for 3.7% of the variance in recall accuracy, whereas pupillary responses at encoding uniquely accounted for 2.0% of the variance in recall accuracy. Approximately 1.1% of the variance in recall accuracy was shared between the two. In other words, of the 6.8% of the variance explained by WMC and mean TEPRs at encoding, 16.18% of the variance was shared between the two constructs while WMC and mean TEPR uniquely contributed 54.41% and 29.41%, respectively. Thus, both constructs accounted for a portion of unique and shared variance and each were needed to account for variation in recall accuracy.

## Discussion

Experiment 1 assessed the extent to which individual differences on a delayed free recall task are due to unique and shared contributions of WMC, the intensity of attention devoted to items at encoding, and encoding strategies. Replicating previous research, results demonstrated that WMC and encoding strategy use both provide unique sources of variance in predicting recall accuracy (Bailey et al., 2008; Unsworth & Spillers, 2010; Unsworth, 2016). People with higher WMC were more likely to recall correct items in comparison to individuals with lower WMC, and people who endorsed the use of more effective encoding strategies were likewise more likely to recall correct items relative to individuals who reported using more ineffective encoding strategies. Inconsistent with this research is the finding that neither self-reported effective (Bailey et al., 2008; Unsworth, 2016) nor ineffective (Unsworth & Spillers, 2010) encoding strategy use were related to WMC. After a closer inspection of the literature, though, it seems likely that differences in paradigms and statistical analyses used among these studies could explain the conflicting findings.[5]

The results of Experiment 1 extended previous research by being the first to demonstrate that the intensity of attention devoted to items at encoding, as indexed by the pupillary response, is an additional factor related to individual differences in recall accuracy performance. TEPRs during encoding were positively related to individual differences in WMC, and this relation continued to strengthen as more to-be-remembered items were introduced during the encoding period (i.e., as serial position increased). TEPRs at encoding also partially accounted for the relation between individual differences in WMC and recall accuracy. Note that since self-reported encoding strategies were unrelated to WMC and the pupillary response, self-reported encoding strategy use cannot explain why high WMC and low WMC individuals in our sample differed in the amount of attention devoted to items at encoding.

Rather, with respect to high WMC individuals, results are most consistent with the *resource allocation hypothesis* (Granholm et al., 1996; Van der Meer et al., 2010). The increase in TEPRs across serial positions for these individuals appears to be due to a cognitive loading function (see Kahneman & Beatty, 1966; Peavler, 1974), such that having more available resources (i.e., larger capacity) supports the ability to engage in more effortful attentional processing until processing load limits are maximized. It, therefore, seems likely that high WMC individuals incorporated each newly presented word into an ongoing strategy, resulting in increased memory load and larger TEPRs (Kahneman & Beatty, 1966)—at least until points of overloading occurred (at which

---

[5] For instance, Bailey et al. (2008) correlated self-reported strategy use on the WMC tasks with the WMC composite score, whereas the present study correlated strategy use on the free recall task with the WMC composite score. Unsworth's (2016) results were based on a latent variable analysis in which strategy reports were obtained across three versions of a delayed free recall paradigm. A closer examination of the zero-order correlations between each WMC measure and self-reported effective strategy use on the delayed free recall task (with a presentation duration of 4 s) reveals substantially weaker correlations.

point TEPRs decline; see Granholm et al., 1996). While it isn't entirely clear whether high WMC individuals' TEPRs were declining or remaining at asymptote between serial positions 8 and 10, prior work suggests little change in pupil dilation when individuals process items at or near capacity limitations (Granholm et al., 1996; Unsworth & Robison, 2015). If high WMC individuals were chunking multiple items together, these individuals could potentially maintain more than 7 items in the current focus of attention (Cowan, 2001). Thus, the resource allocation hypothesis would also explain an asymptote observed throughout these later serial positions.

The results for low WMC individuals, however, are most consistent with the *sustained attention hypothesis* (Unsworth et al., 2009). Since only the negative linear trend reached significance for people with low WMC (quadratic $F < 1$, $p > .10$), there does not appear to be much evidence of loading occurring across serial positions for these individuals. As a result, it seems unlikely that low WMC individuals in Experiment 1 were processing items in the same manner as high WMC individuals. Instead, the finding that WMC related differences in the intensity of attention were not detected until serial position 4 suggests that low WMC individuals devoted the majority of their attention to items early in the encoding period but struggled to maintain such a high level of performance. Note that the *sustained attention hypothesis* does not predict the exact point in which low WMC individuals are no longer able to sustain attention during encoding. It merely proposes that at some point during encoding, people with low WMC will no longer be able to sustain the intensity of intention. As such, the finding that pupil dilation does not decrease for low WMC individuals until serial position 4 does not run contrary to the previously mentioned claim. Collectively, the results of Experiment 1 suggest WMC related differences in the intensity of attention may be attributed to a combination of factors: resource allocation and the ability to consistently sustain attention during encoding.

## Experiment 2

Experiment 2 sought to first replicate the finding that WMC and the intensity of attention to items at encoding are related. If the relation holds, we aimed to further clarify the reason for which low WMC individuals differentially allocate the intensity of attention to items at encoding. Specifically, we implemented a value-directed remembering procedure (see Watkins & Bloom, 1999) to better discern the role of sustained attention at encoding for people with low WMC. Tasks that adopt a value-directed remembering procedure typically ask participants to study lists of words, but words within each list vary in importance. This is achieved by assigning points (aka "values") to-be-remembered words, which are awarded to participants if the words associated with those points are successfully recalled. Participants are instructed to study each word list with the goal of maximizing the number of points earned at recall. Research employing this procedure (Castel, 2008; Castel, Benjamin, Craik, & Watkins, 2002) suggests that participants selectively attend more to high value items than low value items, resulting in superior recall for items deemed more important. In support of this claim, Ariel and Castel (2014) monitored pupil dilation while participants completed a value-directed remembering paradigm, and results revealed larger TEPRs for high value words relative to low value words. Critically, these high value words were also associated with improved recall accuracy. Thus, words labeled as more important received more attention at encoding and were better remembered.

We reasoned that by applying value-directed remembering to a delayed free recall task, we may be able to manipulate encoding conditions in such a way that we can force low WMC individuals to allocate more attention to items presented later in a list (see Stefanidi, Ellis, & Brewer, 2018). Specifically, if (1) words in the delayed free recall task are paired with point values that will be awarded to participants for recalling those items and (2) the goal is to maximize one's recall score, it seems plausible that an ascending word-value order (e.g., 1—shoe, 2—peach, 3—whale) could force low WMC individuals to differentially allocate their attention to mid and recency items rather than primacy items. Of course, it is possible that low WMC individuals will be unable to allocate more attention to mid and recency items than primacy items—a result that would be predicted by the *sustained attention hypothesis*. Encoding requires substantial effort. So, if attentional resources are devoted to encoding the first items in a list and individuals encounter difficulty sustaining such a level of performance, less attention should presumably be devoted to processing successive items (see Healey & Kahana, 2016). Given the notion that low WMC individuals struggle in maintaining high levels of attentional processing, the *sustained attention hypothesis* would, therefore, still predict a decline in TEPRs across serial positions for these individuals regardless of whether or not later serial positions are labeled as more valuable.

If, however, low WMC individuals are able to devote more attention to mid and recency items than primacy items, this result could potentially rule out the sustained attention hypothesis and also suggest that the allocation of the intensity of attention to items at encoding may be strategic in nature (see Ariel & Castel, 2014). Specifically, if low WMC individuals are able to engage in more effortful attentional processing on later items in a list when these items are more important for recall, one may argue that low WMC individuals strategically devote less attention to items that they consider less likely to assist in better recall performance (i.e., they may reserve their resources for items deemed more important). Extending this reasoning to Experiment 1, it's possible that low WMC individuals prioritized primacy items because they were well aware of their resource limitations. In being aware of the fact that remembering most words presented in a list is unlikely, one may decide to prioritize a subset of items—in this case, the first few items presented in a list. Such a finding would seemingly imply that low WMC individuals are not deficient in memory selectivity, the selectivity with which individuals encode important information (see Ariel & Castel, 2014; Watkins & Bloom, 1999). If low WMC individuals strategically devote more attention to items deemed to be most important, we would predict that TEPRs for these individuals should be largest for high value information relative to low value information and either (1) no WMC related differences in memory selectivity or (2) a negative correlation between WMC and memory selectivity. The present study aimed to further evaluate these relations by using a delayed free recall paradigm similar to that reported in Experiment 1.

### Method

#### *Participants*

A total of 134 participants (age range: 18–31 years) were recruited from the University of Oregon's human subject pool. Six participants were excluded from all analyses due to either excessive missing pupil data on the delayed free recall task or because these individuals were not proficient in English (final $N = 128$). Data was collected over two academic quarters, and all participants were compensated with course credit necessary for meeting a course research requirement.

#### *Materials and procedure*

After signing informed consent and providing demographic information, all participants completed the same three shortened measures of WMC used in Experiment 1 (i.e., Ospan, Symspan, and Rspan). Upon completion of the WMC tasks, participants were moved to a dimly lit room where they completed a similar delayed free recall task to that used in Experiment 1, with the exception that the task used here employed a value directed remembering manipulation at encoding. The delayed free recall task was again performed while pupil diameter was simultaneously recorded binocularly at 120 Hz using a Tobii T120 eye-tracker (see Experiment 1 for specifics about task parameters). Participants were also asked to complete two additional measures of long term memory ability upon completion of the delayed free recall task—a paired associates task followed by a picture source recognition task. A secondary question arose as we were analyzing the data from Experiment 1. Namely, how much of the relation between WMC and the intensity of attention is actually due to LTM ability common to various LTM tasks—not just delayed free recall? Since we were interested in exploring the possibility that the relation between WMC and the intensity of attention is actually mediated by individual differences in LTM ability, we sought to compare high ability (high WMC and high LTM) and low ability (low WMC and low LTM) individuals with respect to how the intensity of attention is allocated at encoding. Results pertaining to these analyses are reported in the Appendix. The entire experimental session lasted approximately one and a half hours.

### *WMC tasks*

See Experiment 1.

### *Delayed free recall task with value directed remembering manipulation*

After calibration of the eye-tracker, participants were administered a delayed free recall task similar to the task used in Experiment 1. The primary difference was that each word was now vertically paired with a number. Participants were instructed that the number presented alongside each word represented the importance (or value) of remembering that word (with values ranging from 1 to 10), of which would be awarded to the participant if the accompanying word was correctly recalled. So, if participants studied and recalled “10—Puppy” and “9—Beach” they would receive 10 points for remembering the word “Puppy” and 9 points for remembering the word “Beach”, yielding 19 points total. Participants were then told that the goal of the task was to maximize the amount of points earned for each list. So, remembering words associated with larger values and remembering as many words as possible would be most advantageous. Note that participants were *not* told their accumulated point totals at any point. Participants were presented with a total of 10 lists containing 10 word—value pairs each. Five of the word—value lists appeared in ascending order (e.g., 1—Shoe, 2—Tree, 3—Whale), whereas the other 5 lists appeared in descending order (e.g., 10—Puppy, 9—Beach, 8—Prince). We included a descending word-value order because we reasoned that this particular format would produce similar results to what was observed in Experiment 1. Namely, low WMC individuals should engage in effortful attentional processing primarily among primacy items, whereas high WMC individuals should still engage in effortful strategic processing across primacy and mid list items (resulting in a larger pupillary cognitive load function).

Ascending and descending lists were presented in a blocked format. The order in which participants encountered these lists (i.e., ascending lists first or descending lists first) was counterbalanced across all participants, meaning all 5 ascending lists appeared together followed by all 5 descending lists and vise-versa. Once participants were excluded from analyses based on the abovementioned criteria, 62 participants remained in the ascending first condition and 66 participants were in the descending first condition. Similar to Experiment 1, the task began with a “Ready?” signal onscreen for 2 s, followed by a fixation period lasting 2.5 s (baseline pupil diameter). Each list began with the same “Ready?” signal and fixation period, which were followed by a series of word—value pairs presented individually in the center of the screen for 3 s. Each word—value pair was preceded and followed by a mask of 5 plus signs (e.g., “+ + + + +) for 500 ms. All other task parameters (e.g., the distractor task and recall period) were identical to the delayed free recall task employed in Experiment 1. Proportion of correctly recalled items and mean pupil dilation during the encoding period of each word—value pair were recorded. We also measured memory selectivity for each participant by computing a selectivity index for each participant, which represents an individual’s score relative to a chance score and an ideal score (as described by Castel et al., 2002).[6]

$$\text{Selectivity Index} = \frac{\text{Total Points Earned} - (\text{Chance Score} \times \text{Total Words Recalled})}{\text{Ideal Score} - (\text{Chance Score} \times \text{Total Words Recalled})}$$

### *LTM tasks*

**Paired associates.** Participants were administered 3 lists of 10 word pairs each. All words were common nouns, and each word pair was presented vertically for 2 s. All word pairs were associatively and semantically unrelated. Participants were told that the cue would always be the word on top and the target would be on bottom. After the presentation of the last word participants saw the cue word and ??? in place of the target word. Participants were instructed to type in the target word from the current list that matched cue. Cues were randomly mixed so that the corresponding target words were not recalled in the same order as they were presented. Participants had 5 s to type in the corresponding word. A participant’s score was proportion of items recalled correctly.

**Picture source recognition.** During the encoding phase, participants were presented with a picture (30 total pictures) in one of four different quadrants onscreen for 1 s. Participants were explicitly instructed to pay attention to both the picture (item) as well as the quadrant it was located in (source). At test, participants were presented with 30 old and 30 new pictures in the center of the screen. Participants were required to indicate if the picture was new or if it was old. If the picture was deemed old, they also had to specify what quadrant the picture was presented in via key press. Thus, on each test trial participants pressed one of five keys indicating new, old-top left, old-top right, old-bottom left, or old-bottom right. Participants had 5 s to press the appropriate key to enter their response. A participant’s score was the proportion of correct responses.

## Results

Similar to Experiment 1, WMC was treated as a continuous variable in all analyses by creating a single WMC composite score for each participant via mean standardization of participants’ Ospan, Symspan, and Rspan scores. Note that because no a priori hypotheses were made pertaining to the counterbalancing order factor in Experiment 2 (i.e., ascending lists first vs descending lists first), we did not include this order factor as part of the following ANOVAs. Those results are reported in the Appendix. Descriptive statistics and reliability estimates are listed in Table 5. For correlations between all measures, see Table 6.

### *Behavioral effects*

*Recall accuracy.* First we submitted recall accuracy to a 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) repeated measures ANOVA. Results revealed a main effect of serial position, $F(9, 1143) = 70.55$, $MSE = .05$, $p < .001$, partial $\eta^2 = .36$. Consistent with Experiment 1 and other prior work using delayed free recall (Glanzer & Cunitz, 1966), a clear primacy effect was observed such that words presented at earlier serial positions were better recalled than words presented at later serial

---

[6] For a list ranging in values of 1 through 10, an ideal score for an individual who recalled three words would be 10 + 9 + 8 = 27. A chance score, however, is the average points possible (for a 10 item list, this would be 5.5), which is multiplied by the number of words recalled. Thus, if an individual remembered three words worth 10, 9, and 6 points, that participant’s selectivity index would be (25 − 16.5)/(27 − 16.5) = 0.81.

**Table 5**
Descriptive statistics and reliability estimates for all measures.

| Measure | *N* | *M* | *SD* | Skew | Kurtosis | Reliability |
|---|---|---|---|---|---|---|
| Ospan | 128 | 37.65 | 8.14 | −.79 | .58 | .67 |
| Symspan | 128 | 19.52 | 4.99 | −.75 | .47 | .61 |
| Rspan | 128 | 38.16 | 7.11 | −.49 | .16 | .64 |
| DFR Accuracy | 128 | .53 | .16 | .35 | .10 | .92 |
| Mean TEPR | 128 | .06 | .15 | −.17 | .46 | .93 |
| Mean SI | 128 | .34 | .20 | .07 | .39 | .40 |

*Note.* One person was missing data for the picture source recognition task because they left the experimental session early*;* DFR = delayed free recall; SI = selectivity index.

**Table 6**
Correlations among all measures.

| Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. WMC | – | | | | | | |
| 2. Ospan | $.84^{***}$ | – | | | | | |
| 3. Symspan | $.67^{***}$ | $.32^{***}$ | – | | | | |
| 4. Rspan | $.78^{***}$ | $.61^{***}$ | $.20^{*}$ | – | | | |
| 5. DFR Acc | $.27^{**}$ | $.17^{+}$ | $.18^{*}$ | $.26^{**}$ | – | | |
| 6. Mean TEPR | .10 | .11 | .01 | .12 | $.22^{*}$ | – | |
| 7. Mean SI | −.06 | −.03 | .01 | −.11 | −.05 | −.10 | – |

$^{+}$ $p = .05$.
$^{*}$ $p < .05$.
$^{**}$ $p < .01$.
$^{***}$ $p < .001$.

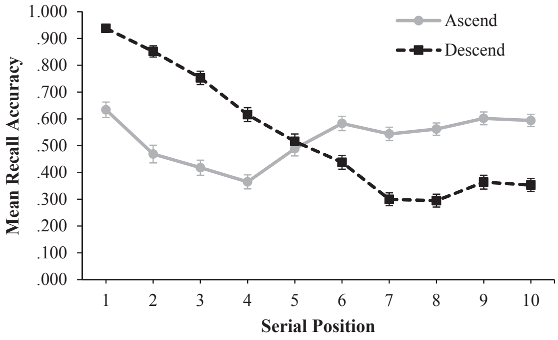

**Fig. 8.** Mean recall accuracy as a function of serial position and list type (ascending vs descending).

positions. Importantly, while there were no main effect of list type ($F = 2.83$, $p = .10$), serial position did interact with list type, $F(9, 1143) = 89.68$, $MSE = .05$, $p < .001$, partial $\eta^2 = .41$. This finding is consistent with Stefanidi et al. (2018), which suggests that our encoding manipulation was successful. Namely, Fig. 8 shows that primacy items were better recalled in descending lists than ascending lists, whereas most mid list and all recency items were better recalled in ascending lists than descending lists. Thus, we are able to force participants to attend more to mid and recency items by labeling these items as more valuable.

To examine whether any of the previously mentioned effects changed as a function of WMC, we next added WMC as a covariate. The 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) repeated measures ANCOVA revealed a main effect of WMC, $F(1, 126) = 9.68$, $MSE = .47$, $p = .002$, partial $\eta^2 = .07$, suggesting higher WMC was associated with superior recall accuracy ($r = .27$, $p = .002$). The only other effect to emerge was a significant interaction between WMC and serial position, $F(9, 1134) = 2.11$, $MSE = .05$, $p = .03$, partial $\eta^2 = .02$. Fig. 9 demonstrates that WMC related differences in recall accuracy were largest for mid-list items. Indeed, the correlation between WMC and mean recall accuracy was strongest for mid-list items ($r = .31$, $p < .001$), followed by primacy items ($r = .19$, $p < .05$), and recency items ($r = .16$, $p = .06$).

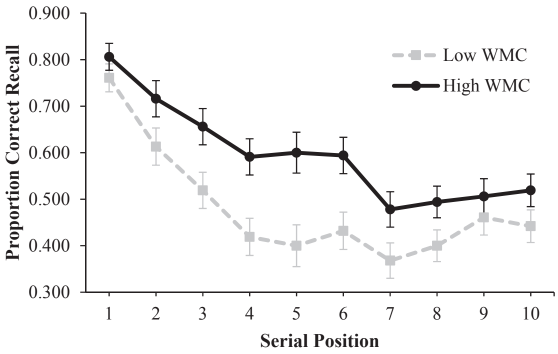

**Fig. 9.** Mean recall accuracy as a function of serial position for high WMC ($n = 32$) and low WMC ($n = 31$) individuals.

*Memory selectivity*. An examination of Table 6 reveals that memory selectivity was not associated with recall accuracy ($r = -.05$, $p > .61$). Nonetheless, to explore potential experimental effects upon memory selectivity, we submitted selectivity index scores to a 2 (list type: ascending vs descending; within-subjects factor) × 5 (list number; within-subjects factor) repeated measures ANOVA. A significant main effect of list type emerged, $F(1, 127) = 243.87$, $MSE = .29$, $p < .001$, partial $\eta^2 = .66$, suggesting participants displayed greater memory selectivity in descending ($M = .57$, $SE = .02$, 95% CI [.54, .61]) than ascending ($M = .10$, $SE = .03$, 95% CI[.05, .15]) lists. There was also a main effect of list number, $F(4, 508) = 2.96$, $MSE = .15$, $p < .05$, partial $\eta^2 = .02$. This finding indicates that memory selectivity increased across lists, positive linear trend: $F(1, 127) = 11.06$, $MSE = .13$, $p = .001$, partial $\eta^2 = .08$. However, the effect of list number was qualified by an interaction with list type, $F(4, 508) = 3.54$, $MSE = .16$, $p < .01$, partial $\eta^2 = .03$, revealing that there was no effect of list number when participants studied descending word-value lists ($F < .34$, $p > .85$). The effect of list number only appeared when participants studied ascending word-value lists, positive linear trend: $F(1, 127) = 17.66$, $MSE = .18$, $p < .001$, partial $\eta^2 = .12$. Thus, while selectivity remained relatively stable across lists when participants studied descending word-value pairs, selectivity increased with continued task experience when participants studied ascending word-value pairs. When adding WMC as a covariate, results revealed no main effect of WMC ($r = -.06$, $F = .44$, $p > .50$) and no experimental interactions with WMC (all *F’s* $< 1$, *p’s* $> .60$).

### *Pupillary effects*

Next, we examined the primary dependent variable of interest, pupillary responses. Again, data from each participant’s left eye was used for analyses, and missing data points were excluded from averaging. Baseline pupil diameter was also calculated via the same method utilized in Experiment 1. Of note, the pupil data for the 3 s encoding phase for each word was similarly broken down into a series of 200 ms timeframes, resulting in 15 total baseline corrected bins. However, bin was not included as a within-subjects factor in the following analyses in order to boost power. Mean pupil dilation for each serial position was calculated by taking the average of all 15 baseline corrected bins.

*Serial position and list type manipulation*. Mean TEPRs during encoding were examined as a function of serial position and list type. The 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) repeated measures ANOVA revealed a main effect of serial position, $F(9, 1143) = 4.49$, $MSE = .01$,

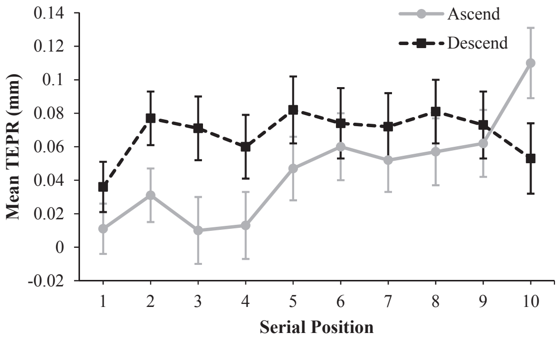

**Fig. 10.** Mean TEPR as a function of serial position and list type (ascending vs descending).

$p < .001$, partial $\eta^2 = .03$. While this result suggests pupil dilation increased across serial positions overall [positive linear trend: $F(1, 127) = 7.87$, $MSE = .05$, $p = .006$, partial $\eta^2 = .06$], serial position significantly interacted with list type, $F(9, 1143) = 6.58$, $MSE = .01$, $p < .001$, partial $\eta^2 = .05$. Fig. 10 demonstrates that pupil dilation did continue to increase across serial positions for ascending lists, positive linear trend: $F(1, 127) = 22.58$, $MSE = .03$, $p < .001$, partial $\eta^2 = .15$. In regards to descending lists, though, TEPRs increased primarily for primacy items, remained at asymptote across mid-list items, and declined across recency items, quadratic trend: $F(1, 127) = 6.46$, $MSE = .04$, $p = .01$, partial $\eta^2 = .05$. These results further support the notion that our encoding manipulation was successful: More attention was devoted to recency items in ascending lists, whereas more attention was allocated to primacy items in descending lists.

Similar to above, we ran another 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) repeated measures ANCOVA with WMC as a covariate, which revealed a different pattern of results than what was observed in Experiment 1. Specifically, there was no main effect of WMC ($F = 1.36$, $p > .24$) on mean TEPRs during encoding ($r = .10$, $p > .24$). However, an interaction between WMC and list type did emerge, $F(1, 126) = 9.80$, $MSE = .15$, $p = .002$, partial $\eta^2 = .07$. Fig. 11 shows that despite high WMC individuals having larger TEPRs across all items in ascending lists ($r = .22$, $p = .01$), both high WMC and low WMC individuals displayed increases in TEPRs across serial positions, positive linear trend: $F(1, 126) = 22.31$, $MSE = .03$, $p < .001$, partial $\eta^2 = .15$. On the other hand, no WMC related differences in pupillary responses emerged on descending lists ($r = -.05$, $p > .58$). And, instead of increasing TEPRs across all serial positions, both high WMC and low WMC individuals now displayed a pattern of results similar to what was observed for high WMC individuals in Experiment 1: an increase in TEPRs across primacy and mid list items, followed by a decrease/plateau in TEPRs across remaining items, quadratic trend: $F(1, 126) = 6.32$, $MSE = .04$, $p = .01$, $\eta^2 = .05$.

**Table 7**
Simultaneous regression predicting recall accuracy.

| Predictor | β | *t* | *sr*$^2$ | *R*$^2$ | *F* |
|---|---|---|---|---|---|
| WMC | .24 | 2.84** | .06 | | |
| Mean TEPR | .19 | 2.21* | .04 | .10 | 7.17** |

* $p < .05$.
** $p < .01$.
*** $p < .001$.

*Regressions.* A simultaneous linear regression model was used to predict delayed free recall accuracy with WMC and mean TEPR as predictors. Results revealed that, together, the predictors significantly accounted for 10.7% of the variance in delayed free recall accuracy, $F(2, 130) = 7.65$, $p = .001$. As demonstrated in Table 7, both WMC ($\beta = .24$, $t(124) = 2.84$, $p < .01$) and mean TEPR($\beta = .19$, $t(124) = 2.21$, $p < .05$) were positively and uniquely related to delayed free recall accuracy. Thus, consistent with Experiment 1, these results suggest delayed free recall performance is partially driven by WMC and the intensity of attention devoted to items during encoding.

## Discussion

Using a combined value-directed remembering/delayed free recall task, Experiment 2 tested the notion that the relationship between WMC and the intensity of attention is partially driven by low WMC individuals' inability to sustain attention throughout encoding. The results revealed a number of findings. First, an examination of our experimental effects suggests that our value directed remembering manipulation was successful in forcing individuals to attend more to either the beginning or end of a list. Participants were more likely to recall recency items in ascending lists and more likely to recall primacy items in descending lists. Moreover, the same can be said for our pupillary effects: In ascending lists, TEPRs were largest for mid list and recency items. On the other hand, descending lists mimicked what was observed in Experiment 1. That is, TEPRs increased across primacy and mid list items but decreased across recency items, meaning primacy TEPRs were largest in descending lists. Thus, consistent with prior research (Ariel & Castel, 2014), high value information was associated with improved recall accuracy and larger TEPRs.

After establishing that our manipulation was effective, we next examined whether the intensity of attention, as indexed by pupillary responses at encoding, was related to fundamental differences in WMC. People with higher WMC again displayed better recall accuracy than people with lower WMC, but the overall correlation between WMC and

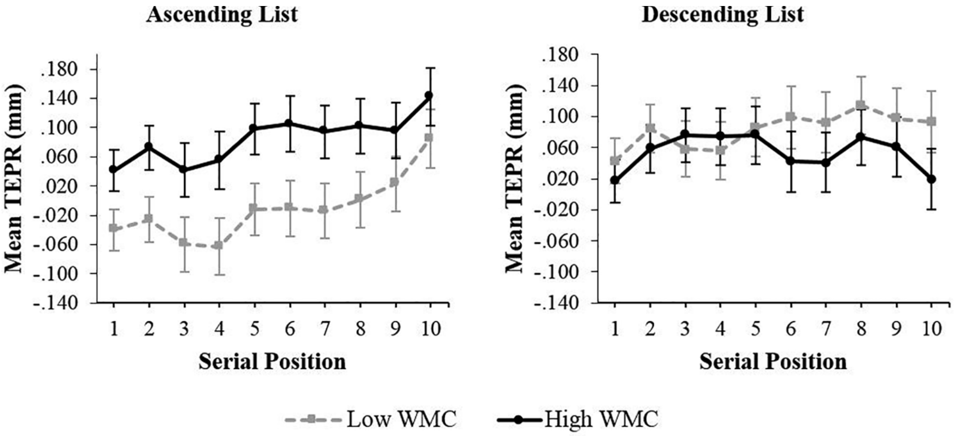

**Fig. 11.** Mean TEPRs as a function of serial position and list type (ascending vs descending) for high WMC ($n = 32$) and low WMC ($n = 31$) individuals.

mean TEPR was not significant ($r = .10$, $p > .24$). Analyses further revealed, though, that despite no main effect of WMC, WMC was related to TEPRs at encoding when participants were presented with ascending lists ($r = .22$, $p = .01$) rather than descending lists ($r = -.05$, $p > .58$).

Critically, while high WMC individuals allocated more attention to all items during encoding of ascending lists, both high WMC and low WMC individuals displayed an increase in TEPRs across serial positions. Because more attention was devoted to mid and recency items than primacy items when participants were presented with an ascending word-value order, the *sustained attention hypothesis* cannot explain why low WMC individuals devote less attention to items during encoding. Presumably, if low WMC individuals struggle to sustain high levels of attentional processing during encoding, low WMC individuals should have been unable to allocate more attention to items presented later in a list (relative to items presented at the beginning of a list).

Instead, the results of Experiment 2 seem to suggest that low WMC individuals are able to engage in effortful attentional processing during encoding among items labeled as more important—just not to the same degree as high WMC individuals. Thus, it appears plausible that given their resource limitations, low WMC individuals may well be aware of the difficult nature of the task. As a result, low WMC individuals may strategically reserve their resources for processing items they believe are more important—at least under conditions of delayed free recall. In support of this claim, the results of Experiment 2 further revealed that low WMC individuals in the present study were not deficient in memory selectivity ($r = -.06$, $p > .50$). While the finding that WMC and memory selectivity were unrelated in the present study is inconsistent with some prior work (Hayes, Kelly, & Smith, 2013; Robison & Unsworth, 2017),[7] we are not alone in detecting no significant relation between WMC and memory selectivity (see Castel, McCabe, & Balota, 2009). Considering research on the relation between WMC and memory selectivity is still largely underway, we don't believe there is reason to doubt that WMC is unrelated to memory selectivity under conditions of delayed free recall.

## General discussion

In two experiments, we examined how the WMC—LTM relationship may be explained by the amount of attention devoted to items at encoding, which was indexed via pupil dilation during the encoding phase of a delayed free recall task. We initially proposed four potential reasons as to why WMC related differences in the intensity of attention may arise (see Fig. 1). The first possibility was the *efficiency hypothesis* (Ahern & Beatty, 1979), according to which low WMC individuals would have to put forth more attentional effort at encoding to achieve a level of recall performance inferior to that of high WMC individuals (i.e., low WMC individuals would have larger TEPRs and worse recall accuracy than high WMC individuals). The second possibility was the *sustained attention hypothesis.* According to this view, low WMC individuals would show a decline in TEPRs across serial positions, whereas high WMC individuals would show no change in TEPRs across serial positions. This pattern of results would arise if low WMC individuals devote less attention to subsequent items at encoding due to their inability to maintain high levels of attentional processing, whereas high WMC individuals' superior attention control abilities would facilitate their ability to consistently sustain attention across the entire encoding period.

The third possibility, the *attentional effort hypothesis* (Heitz, et al., 2008; Van der Meer et al., 2010), was that high WMC individuals would generally devote more attention to items at encoding than low WMC individuals, meaning high WMC individuals would display larger TEPRs across all items. Such a result may arise if high WMC individuals have more attentional resources available for processing items, or if high WMC individuals are more motivated. A final possibility that similarly incorporated attentional resources was the *resource allocation hypothesis* (Granholm et al., 1996; Van der Meer et al., 2010). It could be the case that having more resources may facilitate the ability to employ more effortful encoding strategies until points of overloading occur. That is, individuals may attempt to incorporate each newly presented item during encoding into an ongoing strategy, resulting in increased memory load and larger TEPRs. But, given high WMC individuals have larger maximum capacities, these individuals should be able to maintain more items than low WMC individuals. Hence, a more robust pupillary cognitive load function should be observed for people with high WMC.

The results of Experiment 1 demonstrated that larger TEPRs at encoding were associated with both improved recall accuracy and higher WMC. Regression analyses further revealed that TEPRs at encoding partially accounted for the WMC—LTM relationship. Given the finding that no WMC related differences were observed in self-reported effective or ineffective strategy use, differences in encoding strategies cannot explain why high WMC individuals devoted more attention to items at encoding than their low WMC counterparts. However, a closer examination of TEPRs across serial positions revealed different patterns of results for high WMC and low WMC individuals. People with high WMC displayed an increase in pupil dilation across serial positions, a result predicted by the *resource allocation hypothesis* (Granholm et al., 1996; Van der Meer et al., 2010). On the other hand, people with low WMC showed a decrease in pupil dilation across serial positions, a result best accounted for by the *sustained attention hypothesis*.

Experiment 2 aimed to replicate and extend these results. Specifically, Experiment 2 sought to better elucidate the role of sustained attention at encoding for low WMC individuals. Results revealed that even though higher WMC was again associated with improved delayed free recall accuracy, WMC was related to pupillary responses at encoding only when items presented later in a list were labeled as more important (i.e., when word-value pairs appeared in ascending order and not descending order). Moreover, despite high WMC individuals devoting more attention to all items in this condition, TEPRs for both high WMC and low WMC individuals increased across serial positions. These results contradict the claim from Experiment 1 that low WMC individuals devote less attention to items at encoding because of an inability to sustain attention. If sustained attention were the mechanism, low WMC individuals should not have been able to devote more attention to items presented later in a list.

Instead, the results from Experiments 1 and 2 suggest that low WMC individuals can engage in effortful attentional processing—just not to the same extent as high WMC individuals. As such, WMC related differences in the intensity of attention appear to be due to a combination of the previously mentioned factors: *attentional effort* and *resource allocation.* Considering the lack of support existing for WMC related differences in motivation (Heitz et al., 2008; Robison & Unsworth, 2015, in press; Unsworth & McMillan, 2013), we believe it's likely that given their resource limitations, low WMC individuals may be well aware of the demanding nature of encoding items into LTM—at least under conditions of delayed free recall. Thus, in these situations, low WMC individuals may reserve what little resources they have for processing items deemed most important/advantageous for better task performance. Indeed, in support of this claim, low WMC was not associated with impaired memory selectivity. In fact, the direction of the effect was negative, such that lower WMC was—insignificantly—related with increased memory selectivity. And, applying this logic to Experiment 1, it's possible that in being aware of the fact that remembering most words presented in a list is unlikely, low WMC individuals may have decided to prioritize a subset of items—in this case, primacy items.

---

[7] The lack of an effect in the current study is not entirely surprising given Robison and Unsworth's (2017) effect was marginal at $r = .16$ ($p = .08$). In addition, our methodology substantially differed from the procedures employed by both Hayes et al. (2013) and Robison and Unsworth (2017).

It is important to note, though, that an additional expectation of ours was that a descending word-value order in Experiment 2 would produce similar results to what was observed in Experiment 1. Specifically, we believed that high WMC individuals would continue to engage in effortful strategic processing across primacy and mid list items, resulting in a pupillary cognitive load function, whereas low WMC individuals would primarily engage in effortful attentional processing among primacy items. An examination of the serial position functions for descending lists revealed high WMC individuals in Experiment 2 performed similarly to high WMC individuals in Experiment 1: TEPRs increased across primacy and most mid list items but decreased/plateaued across remaining items. The surprise, then, was that low WMC individuals were seemingly able to match this level of performance.

Admittedly, we do not have a great explanation as to why low WMC individuals were able to equate performance with high WMC individuals on descending lists in Experiment 2. Of course, these results may simply be the product of the inherently messy nature of data. Alternatively, there exists the potential impact of differences in the effects of instructions on these discrepant results. In Experiment 1, participants were given no instructions and relied on their own, self-generated strategies for encoding items. Conversely, in Experiment 2, participants were explicitly told to pay more attention to high-value items. In telling participants what items are most important to remember, we may have minimized variation in encoding strategies and subsequently altered the way in which certain individuals process items at encoding. For instance, perhaps the cognitive loading function wasn't as robust in descending lists for high WMC individuals in Experiment 2 because they now elected not to incorporate recency items (i.e., low value items) into their ongoing strategies. Indeed, when examining Experiment 2's descending lists, the proportion of correctly recalled recency items for high WMC individuals was .36 ($SD$ = .22)—a level of performance significantly lower than what was observed for high WMC individuals in Experiment 1 ($M$ = .48, $SD$ = .21; $t(67) = 2.28$, $p = .026$). These results suggest that high WMC individuals in Experiment 2 ignored low value information (consistent with Robison & Unsworth, 2017) and, instead, put more effort into remembering primacy items (Experiment 1 $M$ = .81, Experiment 1 $SD$ = .13; Experiment 2 $M$ = .88, Experiment 2 $SD$ = .15). This difference among primacy items, however, was marginal, $t(67) = 1.87$, $p = .066$.

Taken altogether, it appears that WMC is related to the intensity of attention at encoding, but this relation is small. Thus, while the intensity of attention is one source of variation explaining why high WMC individuals outperform low WMC individuals in delayed free accuracy, the relation between WMC and LTM is largely driven by other factors—such as search efficiency (Miller & Unsworth, in press; Unsworth & Engle, 2007; Unsworth, 2007) and variation in monitoring abilities (Unsworth & Brewer, 2010). Nonetheless, the current study contributes to the literature by demonstrating that WMC related differences in pupillary responses at encoding appear to be driven by the amount of resources available for processing items, as well as the strategic allocation of those resources. Given their excess capacity, high WMC individuals appear to process the entire word list as a single entity. That is, high WMC individuals may incorporate each subsequent word into an ongoing strategy. On the other hand, people with low WMC seemingly compensate for a lack of available resources by selectively focusing their attention on what they deem to be the most valuable items (see also Middlebrooks, Kerr, & Castel, 2017). These results suggest that attentional processes operating at encoding must be taken into consideration when trying to better elucidate reasons for which some individuals (e.g., people with high WMC) are better able to recall items from LTM than others.

## Appendix A

### A.1. Counterbalancing analyses

Given no a priori hypotheses were made pertaining to the counterbalancing factor in Experiment 2 (i.e., ascending lists first vs descending lists first), we elected not to include this factor as part of the ANOVAs within the manuscript. These results are reported here. Note that the patterns of effects described below remained unchanged when adding WMC as a covariate (no additional interactions were observed). To examine possible interactions between WMC and condition, we added WMC as a between subjects factor, instead of a covariate. Importantly, all interactions between WMC and condition were non-significant; all $p$'s $> .15$, $F$'s $< 1.33$).

### A.2. Behavioral effects

*Recall Accuracy.* A 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) × 2 (condition: ascending-lists-first vs descending-lists-first; between-subjects factor) repeated measures ANOVA revealed no main effect of condition on recall accuracy ($F < 1, p > .44$). There was, however, a significant interaction between list type and condition, $F(1, 126) = 21.51$, $MSE = .07$, $p < .001$, partial $\eta^2 = .15$, indicating more items were recalled in ascending lists when participants were in the ascending-lists-first condition ($M$ = .56, $SE$ = .02) than the descending-lists-first condition ($M$ = .49, $SE$ = .02). There was also a significant interaction between serial position and condition, $F(9, 1134) = 4.32$, $MSE = .05$, $p < .001$, partial $\eta^2 = .03$, suggesting more primacy items were recalled in the ascending-lists-first condition than the descending-lists-first condition. However, both of these interactions were qualified by a significant three-way interaction between serial position, list type, and condition, $F(9, 1134) = 9.64$, $MSE = .05$, $p < .001$, partial $\eta^2 = .07$. The three-way interaction suggests that while condition did not influence the size of the primacy effect in descending lists, the size of the primacy effect changed as a function of condition in ascending lists. Specifically, the primacy effect in ascending lists was larger when participants were shown ascending lists first rather than descending lists first.

*Memory selectivity.* We submitted selectivity index scores to a 2 (list type: ascending vs descending; within-subjects factor) × 5 (list number; within-subjects factor) × 2 (condition: ascending-lists-first vs descending-lists-first; between-subjects factor) repeated measures ANOVA. Results revealed a significant main effect of condition, $F(1, 126) = 8.93$, $MSE = .36$, $p < .01$, partial $\eta^2 = .07$, indicating memory selectivity was greater when participants received descending lists first ($M$ = .39, $SE$ = .02) rather than ascending lists first ($M$ = .29, $SE$ = .02). An interaction between list type and condition also emerged, $F(1, 126) = 21.41$, $MSE = .25$, $p < .001$, partial $\eta^2 = .15$. The list type × condition interaction suggests that condition (i.e., whether participants first studied descending or ascending lists) had no effect on memory selectivity when participants studied descending lists. Instead, memory selectivity on ascending lists was higher when ascending lists were presented last ($M$ = .21, $SE$ = .04, 95% CI [.14, .28]) rather than when ascending lists were presented first ($M = -.02$, $SE$ = .04, 95% CI [−.09, .06]). All other effects were non-significant ($F$'s $< 1.92$, $p$'s $> .10$).

### *A.3. Pupillary effects*

*Serial position and list type manipulation.* The 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) × 2 (condition: ascending-lists-first vs descending-lists-first; between-subjects factor) repeated measures ANOVA revealed a main effect of condition, $F(1, 126) = 6.73$, $MSE = .44$, $p = .01$, partial $\eta^2 = .05$, indicating pupil dilation was larger, in general, when ascending lists were encountered first ($M = .09$, $SE = .02$), as opposed to when descending lists were encountered first ($M = .03$, $SE = .02$). A three-way interaction between serial position, list type, and condition also reached significance, $F(9, 1134) = 2.13$, $MSE = .01$, $p < .05$, partial $\eta^2 = .02$. The three-way interaction suggests that TEPR differences (across serial positions) between ascending and descending lists were larger when descending lists were presented before ascending lists.

### *A.4. Individual differences in LTM*

To further investigate individual differences in LTM abilities, we computed a single LTM composite score for each participant by averaging participants' standardized scores on the paired associates and picture source recognition tasks. The LTM composite score was likewise treated as a continuous variable in all analyses, meaning LTM was used as a categorical variable for graphical purposes only (i.e., top 25% of performers on the LTM tasks were classified as high LTM ability individuals, and the lowest 25% performers were considered low LTM ability individuals). Descriptive statistics and reliability estimates are listed in Table A.1. For correlations between all measures, see Table A.2.

**Table A1**
Descriptive statistics and reliability estimates for all measures.

| Measure | *N* | *M* | *SD* | Skew | Kurtosis | Reliability |
|---|---|---|---|---|---|---|
| Ospan | 128 | 37.65 | 8.14 | −.79 | .58 | .67 |
| Symspan | 128 | 19.52 | 4.99 | −.75 | .47 | .61 |
| Rspan | 128 | 38.16 | 7.11 | −.49 | .16 | .64 |
| DFR Accuracy | 128 | .53 | .16 | .35 | .10 | .92 |
| PA Accuracy | 128 | .41 | .26 | .37 | −.84 | .84 |
| PicSource Accuracy | 127 | .72 | .18 | −1.39 | 2.15 | .93 |
| Mean TEPR | 128 | .06 | .15 | −.17 | .46 | .93 |
| Mean SI | 128 | .34 | .20 | .07 | .39 | .40 |

*Note.* One person was missing data for the picture source recognition task because they left the experimental session early*;* DFR = delayed free recall; SI = selectivity index; PA = paired associates; PicSource = picture source recognition.

**Table A2**
Correlations among all measures.

| Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. WMC | – | | | | | | | | | |
| 2. Ospan | .84*** | – | | | | | | | | |
| 3. Symspan | .67*** | .32*** | – | | | | | | | |
| 4. Rspan | .78*** | .61*** | .20* | – | | | | | | |
| 5. DFR Acc | .27** | .17+ | .18* | .26** | – | | | | | |
| 6. Mean TEPR | .10 | .11 | .01 | .12 | .22* | – | | | | |
| 7. Mean SI | −.06 | −.03 | .01 | −.11 | −.05 | −.10 | – | | | |
| 8. LTM | .35*** | .28** | .21* | .32*** | .62*** | .19* | −.10 | – | | |
| 9. PA Acc | .26** | .21* | .10 | .27** | .68*** | .14 | −.14 | .85*** | – | |
| 10. PicSource Acc | .35*** | .27** | .26** | .28** | .36*** | .18* | −.04 | .84*** | .43*** | – |

+ $p = .05$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

*Regressions*. First, a hierarchical linear regression model was used to examine whether a composite score reflecting LTM abilities independent of delayed free recall account for any of the shared variance between WMC and the intensity of attention, as indexed by mean TEPR. Results indicated that WMC, mean TEPR, and LTM jointly explained 39% of the variability in delayed free recall accuracy, $F(3, 123) = 26.33$, $p < .001$. Thus, the full model (i.e., Step 2) adding LTM as a predictor significantly accounted for variance in recall accuracy above and beyond what was explained by the reduced model in Step 1, $\Delta R^2 = .29$, $F(1, 123) = 58.04$, $p < .001$. The full model revealed a significant, positive effect of LTM ability, $\beta = .58$, $t(123) = 7.62$, $p < .001$. Unsurprisingly, higher scores on the composite LTM variable were associated with better recall accuracy on the delayed free recall task. However, with LTM added as a predictor, WMC and mean TEPR no longer accounted for any unique variance in delayed free recall accuracy (both $p$'s $> .17$; see Table A.3). Hence it appears that these factors no longer predict recall accuracy because of their shared variance with episodic memory abilities independent of the task at hand.

To further examine the possibility that LTM ability mediates the WMC—recall relation, as well as the intensity of attention—recall relation, we next used variance partitioning methods (see Chuah & Maybery, 1999; Cowan et al., 2005; Unsworth et al., 2014) to examine shared and unique contributions of WMC, mean TEPR, and LTM ability in predicting delayed free recall accuracy. Similar to Experiment 1, we obtained $R^2$ values for each predictor via a series of simultaneous regression analyses. For each variable entering the regression, the zero-order correlations from Table A.2 were used.

**Table A3**
Stepwise regression predicting recall accuracy for Experiment 2.

| Predictor | β | *t* | $sr^2$ | $R^2$ | *F* |
|---|---|---|---|---|---|
| Step 1 | | | | | |
| WMC | .24 | 2.84** | .06 | | |
| Mean TEPR | .19 | 2.21* | .04 | .10 | 7.17** |
| Step 2 | | | | | |
| WMC | .05 | .61 | .00 | | |
| Mean TEPR | .10 | 1.37 | .01 | | |
| LTM | .58 | 7.62*** | .29 | .39 | 26.33*** |

* $p < .05$.
** $p < .01$.
*** $p < .001$.

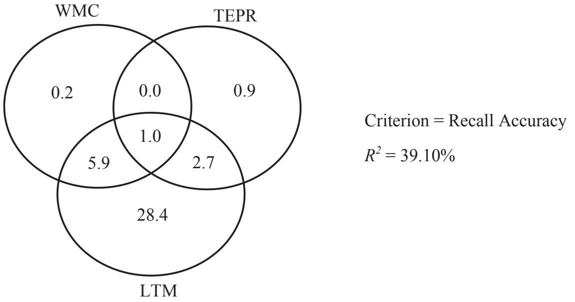

**Fig. A1.** Venn diagrams representing the shared and unique variance between WMC, mean TEPR, and LTM in predicting DFR recall accuracy.

Fig. A.1 reveals that, altogether, WMC, mean TEPR, and LTM ability explained 39.10% of the variance in delayed free recall accuracy. Of this variance, only 1% was shared by all three constructs, whereas the remaining 38.1% was attributed to both unique and shared variance across the three constructs. Namely, WMC and mean TEPR contributed little unique variance (0.2% and 0.9%, respectively) in recall accuracy, whereas LTM ability accounted for a considerable amount of unique variance (28.4%). As such, most of the variance explained by WMC (5.9%) and mean TEPR (2.7%) was shared with LTM ability, and LTM ability accounted for all of the shared variance between WMC and the intensity of attention. Collectively, these results suggest that people who devote more attention to items during encoding on delayed free recall tasks likely devote more attention to items during encoding on other LTM tasks, such as paired associates and picture-source recognition tasks. And, despite the shared variance between WMC and mean TEPR being largely driven by LTM ability, this relation between pupillary responses at encoding and LTM ability does not appear to be driven by individual differences in WMC. Thus, all three constructs need to be taken into consideration when trying to explain variation in recall accuracy.

Given the finding that LTM ability fully accounted for the relation between WMC and mean TEPR when predicting delayed free recall accuracy, we next examined whether LTM related differences in pupillary responses at encoding would be similar to WMC related differences. We ran the same pupillary analyses described previously but added LTM as a covariate, instead of WMC. Since adding LTM as a covariate resulted in no changes with respect to the aforementioned experimental effects, only effects associated with LTM are reported.

### *A.5. Pupillary effects*

*Serial position and list type manipulation.* In the following analysis we submitted mean TEPRs to a 2 (list type: ascending vs descending; within-subjects factor) × 10 (serial position; within-subjects factor) repeated measures ANCOVA with LTM as a covariate. Critically, a different pattern of pupillary results arose with respect to individual differences in LTM abilities (relative to WMC). Namely, there was a main effect of LTM, $F(1,125) = 4.69$, $MSE = .45$, $p = .03$, partial $\eta^2 = .04$, suggesting better LTM ability was associated with larger TEPRs during encoding, in general ($r = .19, p < .05$). Moreover, LTM did not interact with list type ($F < 1; p > .49$). Instead, LTM interacted with serial position, $F(9,1125) = 3.74$, $MSE = .01, p < .001$, partial $\eta^2 = .03$. Fig. A.2 shows that following serial position 1, TEPRs continued to increase across the next few items, reach asymptote across mid list items, and decline across recency items for high LTM individuals [quadratic trend: $F(1,30) = 13.92$, $MSE = .03, p = .001$, partial $\eta^2 = .32$], whereas TERPs slightly declined across primacy items and subsequently increased across all remaining items for low LTM individuals [linear trend: $F(1,31) = 4.14$, $MSE = .05$, $p = .051$, partial $\eta^2 = .12$; quadratic trend: $F(1,31) = 3.99$, $MSE = .02$, $p = .055$, partial $\eta^2 = .11$]. Therefore, despite no differences in the amount of attention devoted to serial position 1 ($r = .10$, $p = .27$), substantial LTM related differences in TEPRs emerged across mid list items ($r = .24$, $p = .007$) but disappeared across recency items ($r = .07$, $p = .41$).

In summary, variance partitioning methods revealed that the shared variance between WMC and TEPRs at encoding was fully mediated by LTM ability, and, notably, pupil dilation at encoding more strongly related to individual differences in LTM than individual differences in WMC. Higher LTM ability was related to overall larger TEPRs at encoding, and LTM related differences in pupillary responses emerged across serial positions, irrespective of list type. More specifically, TEPRs increased across primacy and mid list items, reached an asymptote, and soon began to decrease across the last few items—a result largely consistent with the pattern of results observed for high WMC individuals in both Experiments. Low LTM individuals, however, displayed a different pattern: TERPs slightly declined across primacy items and then increased across all remaining items. This result is similar to what was observed for low WMC individuals on ascending lists in Experiment 2, but inconsistent with how low WMC individuals

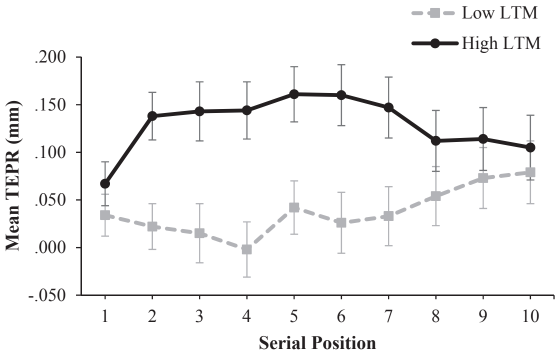

**Fig. A2.** Mean TEPR as a function of serial position for high LTM ($n$ = 31) and low LTM ($n$ = 32) individuals.

performed on descending lists in Experiment 2, as well as Experiment 1. Collectively, these results suggest that while LTM abilities play an important role in explaining WMC related differences in the intensity of attention, important dissociations exist between these constructs. Therefore, it is important for researchers to differentiate between individual differences in WMC and LTM in future work.

## References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General, 131*, 567–589.

Ahern, A., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science, 205*, 1289–1292.

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology and Aging, 13*, 405–423.

Ariel, R., & Castel, A. D. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental Brain Research, 232*, 337–344.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403–450.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Vol. Ed.), *The psychology of learning and motivation: Vol. 8*, (pp. 47–89). New York: Academic Press.

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*, 518–540.

Bailey, H., Dunlosky, J., & Hertzog, C. (2009). Does differential strategy use account for age-related deficits in working memory performance? *Psychology and Aging, 24*, 82–92.

Bailey, H., Dunlosky, J., & Kane, M. J. (2008). Why does working memory span predict complex cognition? Testing the strategy affordance hypothesis. *Memory & Cognition, 36*, 1383–1390.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.). *Handbook of psychophysiology* (pp. 142–162). New York: Cambridge University Press.

Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye opening: Pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues. *Psychological Science, 20*, 1313–1315.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *The Journal of Neuroscience, 33*, 2199–2204.

Braem, S., Coenen, E., Bombeke, K., van Bochove, M. E., & Notebaert, W. (2015). Open your eyes for prediction errors. *Cognitive, Affective, & Behavioral Neuroscience, 15*, 374–380.

Castel, A. D. (2008). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. In A. S. Benjamin, & B. H. Ross (Vol. Eds.), *The psychology of learning and motivation: Vol. 48*, (pp. 225–270). London: Academic Press.

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition, 30*, 1078–1085.

Castel, A. D., McCabe, D. P., & Balota, D. A. (2009). Memory efficiency and the strategic control of attention at encoding: Impairments of value-directed remembering in Alzheimer's disease. *Neuropsychology, 23*, 297–306.

Chuah, Y. M. L., & Maybery, M. T. (1999). Verbal and spatial short-term memory: Common sources of developmental change? *Journal of Experimental Child Psychology, 73*, 7–44.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 97–185.

Cowan, N., Elliot, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42–100.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.

Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition, 29*, 247–253.

Dunlosky, J., & Kane, M. J. (2007). The contribution of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology, 60*, 1227–1245.

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience, 16*, 1146–1153.

Engle, R. W. (1975). Pupillary measurement and release from proactive inhibition. *Perceptual and Motor Skills, 41*, 835–842.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Vol. Ed.), *The psychology of learning and motivation: Vol. 44*, (pp. 145–199). NY: Elsevier.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence. A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309–331.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto word pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavioral Research Methods & Instrumentation, 14*, 375–399.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 252–269.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351–360.

Goldinger, S. D., He, Y., & Papesh, M. H. (2009). Deficits in cross-race face learning: Insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 35*, 1105–1122.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science, 21*, 90–95.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*, 457–461.

Granholm, E., Morris, S. K., Sarkin, A. J., Asarnow, R. F., & Jeste, D. V. (1997). Pupillary responses index overload of working memory resources in schizophrenia. *Journal of Abnormal Psychology, 106*, 458–467.

Hayes, M. G., Kelly, A. J., & Smith, A. D. (2013). Working memory and the strategic control of attention in older and younger adults. *The Journals of Gerontology: Series B, 68*, 176–183.

Healey, M. K., & Kahana, M. J. (2016). A four component model of age-related memory change. *Psychological Review, 123*, 23–69.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology, 45*, 119–129.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem solving. *Science, 143*, 1190–1192.

Janisse, M. P. (1977). *Pupillometry: The psychology of the pupillary response.* Washington, D.C.: Hemisphere Publishing Co.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology, 47*, 310–339.

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly Journal of Experimental Psychology, 64*, 1971–1989.

Kahneman, D. (1973). *Attention and effort.* Upper Saddle River, NJ: Prentice Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583–1585.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 336–358.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence, 14*, 389–433.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition, 29*, 10–17.

Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively distracted: Divided attention and memory for important information. *Psychological Science, 28*, 1103–1115.

Miller, A. L. and Unsworth, N. (in press). Individual differences in working memory capacity and search efficiency. Memory & Cognition.

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology, 48*, 1532–1543.

Papesh, M. H., Goldinger, S. D., & Hout, M. C. (2012). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology, 83*, 56–64.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology, 11*, 559–566.

Phaf, R. H., & Wolters, G. (1993). Attentional shifts in maintenance rehearsal. *American Journal of Psychology, 106*, 353–382.

Robison, M. K. and Unsworth, N. (in press). Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Robison, M. K., & Unsworth, N. (2015). Working memory capacity offers resistance to mind-wandering and external distraction in a context specific manner. *Applied Cognitive Psychology, 29*, 680–690.

Robison, M. K., & Unsworth, N. (2017). Working memory capacity, strategic allocation of study time, and value-directed remembering. *Journal of Memory and Language, 93*, 231–244.

Rohrer, D. (1996). On the relative and absolute strength of a memory trace. *Memory & Cognition, 24*, 188–201.

Rohrer, D., & Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition, 22*, 511–524.

Rowland, D. C., & Kentros, C. G. (2008). Potential anatomical basis for attentional modulation of hippocampal neurons. *Annals of the New York Academy of Sciences, 1129*, 213–224.

Samuels, E. R., & Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function. Part I: Principles of functional organization. *Current Neuropharmacology, 6*, 235–253.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience, 10*, 211–223.

Stefanidi, A., Ellis, D. M., & Brewer, G. A. (2018). Free recall dynamics in value-directed remembering. *Journal of Memory and Language, 100*, 18–31.

Sterpenich, V., D'Argembeau, A., Desseilles, M., Balteau, E., Albouy, G., Vandewalle, G., ... Maquet, P. (2006). The locus coeruleus is involved in the successful retrieval of emotional memories in humans. *Journal of Neuroscience, 26*, 7416–7423.

Tsukahara, J. S., Harrison, T. L., & Engle, R. W. (2016). The relationship between baseline pupil size and intelligence. *Cognitive Psychology, 91*, 109–123.

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49*, 446–468.

Unsworth, N. (2007). Individual differences in working memory capacity and episodic memory: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1020–1034.

Unsworth, N. (2009b). Examining variation in working memory capacity and retrieval in cued recall. *Memory, 17*, 386–396.

Unsworth, N. (2009a). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition, 37*, 837–849.

Unsworth, N. (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable analysis. *Acta Psychologica, 134*, 16–28.

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 50–61.

Unsworth, N., & Brewer, G. A. (2010). Variation in working memory capacity and intrusions: Differences in generation or editing? *European Journal of Cognitive Psychology, 22*, 990–1000.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There's more to working memory capacity-fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review, 16*, 931–937.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104–132.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory. *Cognitive Psychology, 71*, 1–26.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498–505.

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 39*, 832–842.

Unsworth, N., Miller, J. D., Lakey, C. E., Young, D. L., Meeks, J. T., Campbell, W. K., ... Goodie, A. S. (2009). Exploring the relations among executive functions, fluid intelligence, and personality. *Journal of Individual Differences, 30*, 194–200.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent variable analysis of the relationship between processing and storage. *Memory, 17*, 635–654.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review, 22*, 757–765.

Unsworth, N., & Robison, M. K. (2017b). The importance of arousal for variation in working memory capacity and attention control: A latent variable pupillometry study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1962–1987.

Unsworth, N., & Robison, M. K. (2017a). A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin & Review, 24*, 1282–1311.

Unsworth, N., & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual-retrieval. *Psychonomic Bulletin & Review, 17*, 200–205.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2010). The contributions of primary and secondary memory to working memory capacity: An individual differences analysis of immediate free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 240–247.

Van Der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., ... Wartenburger, I. (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology, 47*, 158–169.

Van Gerven, P. W. M., Paas, F., Van Merrienboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41*, 167–174.

Watkins, M. J., & Bloom, L. C. (1999). Selectivity in memory: An exploration of willful control over the remembering process. Unpublished manuscript.