

THESIS

THE NATURE OF PROCESSING IN WORKING MEMORY: TEMPORAL-CONTEXTUAL  
CUES AND CHARACTERISTICS

Submitted by

Vanessa Maria Loaiza-Kois

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2010

COLORADO STATE UNIVERSITY

February 5, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY VANESSA MARIA LOAIZA-KOIS ENTITLED THE NATURE OF PROCESSING IN WORKING MEMORY: TEMPORAL-CONTEXTUAL CUES AND CHARACTERISTICS BE ACCEPTED AS FULFILLING IN PART REQUIRMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate work



Brian L. Tracy



Matthew G. Rhodes



Advisor, David P. McCabe



Department Chair, Ernest L. Chavez

## ABSTRACT OF THESIS

### THE NATURE OF PROCESSING IN WORKING MEMORY: TEMPORAL-CONTEXTUAL CUES AND CHARACTERISTICS

The primary-secondary memory hypothesis proposes that processing in working memory requires maintaining activated representations in a capacity-limited primary memory while re-accessing representations from secondary memory that have been displaced when the limits of primary memory have been exceeded. An implication is that simple span list lengths that exceed primary memory involve the same temporal-contextual search of secondary memory that is utilized in all trials of a complex span task. A series of experiments tested whether a) temporal-contextual cues could successfully elicit items that were studied during operation span and supra-span trials of word span as opposed to sub-span word span trials, and b) whether temporal-contextual characteristics are more phenomenologically memorable for items studied in trials of operation span and supra-span trials of word span than for sub-span trials of word span. Temporal-contextual cues and characteristics, however, were more accessible for operation span items than for items from simple span at any list length. Implications are discussed in light of this recent theory of working memory capacity as well as models of temporal distinctiveness.

Vanessa Maria Loaiza-Kois  
Department of Psychology  
Colorado State University  
Fort Collins, CO 80523  
Spring 2010

## ACKNOWLEDGEMENTS

I would like to thank my committee for their helpful comments and direction for this paper. I would also like to thank my undergraduate assistants who helped me to collect and enter these data. Finally, I would especially like to thank my adviser, Dr. David McCabe, for his guidance and support, without which this project would not have been possible.

DEDICATION

For my family.

## TABLE OF CONTENTS

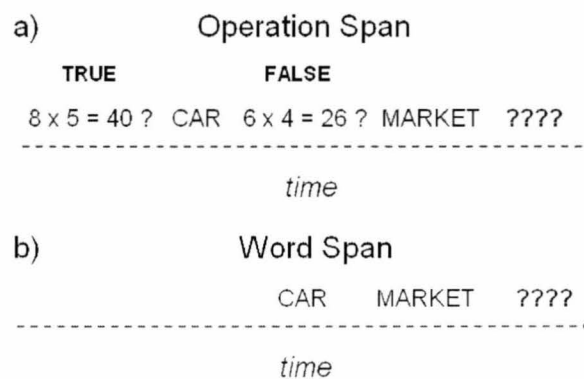
Section	Page Number
Introduction	1
Experiment 1	9
Experiment 2	19
Experiment 3	24
Experiment 4	31
General Discussion	36
References	45

The dynamic memory that takes place from moment to moment has fascinated researchers since the first laboratories of psychology were created (Ebbinghaus, 1885/1913; Wundt, 1897/1969a; Wundt, 1912/1973). What are the processes that underlie the ability to maintain such transient information, and by what means is that information transferred to and retrieved from long term memory? Furthermore, why is some information better remembered than other information? These fundamental questions of memory have been frequently studied, recently with respect to the processes that are involved in working memory.

Working memory traditionally has been conceptualized as a limited capacity system that allows people to briefly maintain representations while simultaneously engaging in other processing activities (Baddeley & Hitch, 1974). The traditional model includes two slave systems, the phonological loop and the visual spatial sketchpad, which are responsible for the passive storage or maintenance of modality-specific information. Additionally, the central executive is considered to be responsible for the allocation of resources between maintaining and processing information.

Complex span tasks tax the central executive by measuring the degree to which a person can maintain and manipulate information effectively. This measure of efficiency is referred to as working memory capacity. Complex span tasks measure working memory capacity by including both a processing component and a storage or maintenance component. For example, the operation span task (Turner & Engle, 1989) requires the participant to solve a multiplication problem, and also try to maintain a word that is presented subsequent to the math problem. This problem-word sequence is typically repeated two to five times per trial until a cue is given to the participant to recall all of the words from that trial out loud (see Figure 1a). Thus, the operation span taps central executive processing in requiring both a cognitively-demanding processing

component (i.e., the multiplication problem) as well as temporary maintenance of information (i.e., recalling the words from the trial). Conversely, simple span tasks are thought to measure the capacity of the so-called “slave systems,” which only require brief storage of information without requiring any additional central executive resources being engaged. For example, the word span task presents several successive words, usually between two and seven, followed by a cue to recall the words of that trial (see Figure 1b). Arguably, what distinguishes complex span tasks and simple span tasks is the degree to which the cognitive demands of the task engage the central executive.



*Figure 1.* Examples of (a) a complex span task (operation span) and (b) a simple span task (word span).

In addition to its theoretical importance, the central executive component of the working memory model has been proposed as especially vital in predicting higher order cognitive abilities, such as general fluid intelligence (Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004) and reading comprehension (Daneman & Carpenter, 1980; Turner & Engle, 1989). Specifically, evidence indicates that complex span tasks are more predictive of higher cognitive abilities than are simple span tasks. This may be due to the important role of the central executive in allocating resources between the different cognitive demands of complex span tasks.

*Primary-Secondary Memory Hypothesis*



Recently, an alternative conceptualization of the relation between simple and complex span tasks and different memory processes was examined. According to the primary-secondary memory hypothesis (Unsworth & Engle, 2007a), simple and complex span tasks involve retrieval from both primary and secondary memory. Primary memory involves retrieval of information that can be maintained briefly in the focus of attention, whereas secondary memory involves retrieval from long term memory (Watkins, 1974).

Unsworth and Engle (2006a) found that performance on simple span tasks resembles performance of complex span tasks when the simple span tasks are more difficult in nature. Specifically, simple span trials of longer list lengths are more strongly correlated with higher fluid intelligence than simple span trials of shorter list lengths (Unsworth & Engle, 2006a; Unsworth & Engle, 2007b), and these same supra-span list lengths also load on a different factor than sub-span list lengths of the same simple span task (Unsworth & Engle, 2006a). Thus, simple span trials of longer list lengths appear to measure the same underlying construct as complex span tasks. The importance of secondary memory in predicting higher-ordered cognitive abilities has been found in children as well (De Alwis, Myerson, Hershey, & Hale, 2009). This suggests that the distinctions commonly made between simple and complex span tasks are not as straightforward as originally thought. Instead, Unsworth and Engle have argued that both simple and complex span tasks measure the same mechanisms, but to different degrees. Complex span tasks require retrieval from primary and secondary memory on each trial, whereas simple span tasks of shorter list lengths only require retrieval from primary memory. Simple span tasks of longer list lengths require retrieval from both primary and secondary memory. Hence, a cue-dependent search process of secondary memory is necessary during complex span tasks and simple span trials of longer list lengths in order to retrieve the items that have been displaced from primary memory (Unsworth & Engle, 2007b).

The hypothesis that span task performance is a combination of temporary maintenance in primary memory and a cue-dependent search of secondary memory has been further investigated

in performance on immediate and delayed tests. McCabe (2008) found that performance differed at immediate and delayed recall for both simple span and complex span tasks. Specifically, delayed recall performance was greater for items that had originally been processed during a complex span task compared with a simple span task, whereas the opposite occurred for immediate recall performance (McCabe, 2008). This cross-over interaction has been explained according to the covert retrieval hypothesis, which suggests that repeated retrieval from secondary memory during complex span tasks in turn creates stronger long term memory cues for those items at delayed recall (see Figure 2). Because the simple span task trials in these studies only included two to four words to recall at each trial, those items should remain within primary memory without being displaced (which is also consistent with Cowan's (1993) focus of attention model). Hence, no controlled search of secondary memory is necessary for sub-span list lengths of a simple span task, whereas complex span tasks necessitate covert retrieval strategies that in turn promote stronger retrieval cues that improve performance of those items at delayed recall (McCabe, 2008). The covert retrieval hypothesis was further illustrated in that the items from a complex span task that are most likely to be recalled at delay are the items at the beginning of the span task trials, resulting in a linear downward trend in proportion of items that are recalled at later serial positions. These data are similar to the robust negative recency effect ( Craik, 1970; Craik, Watkins, & Gardiner, 1970), in which the items at the end of a list that are most likely to be recalled on an immediate free recall test are the least likely to be recalled at a final delayed recall test. Furthermore, the effect of serial position was not present for delayed recall of simple span items (McCabe, 2008). Taken together, these data suggest both that earlier-presented items have more opportunities to be covertly retrieved throughout a complex span trial, and that covert retrieval strategies are unnecessary for items that are able to be maintained within primary memory. Thus, retrieval from secondary memory appears to be important for successful performance on tasks in which items are displaced from primary memory, and/or require central executive resources.

## Operation Span

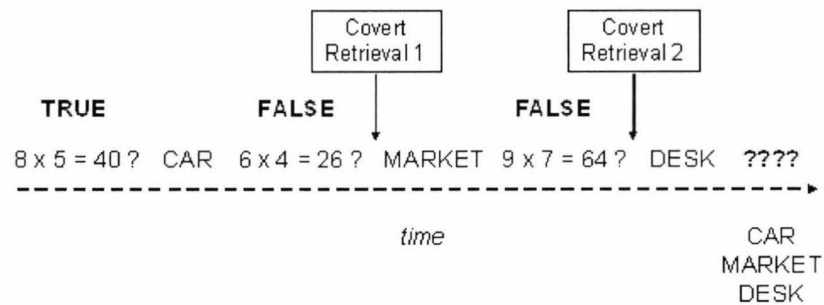


Figure 2. The covert retrieval model of complex span task performance.

### *Processes Engaged by Primary and Secondary Memory Requirements*

As mentioned previously, a simple span task can require retrieval from secondary memory if performance requires retrieval of items that have been displaced from primary memory (Unsworth & Engle, 2007a). Hence, there may be a fundamental difference in the type of processing at encoding between tasks that require retrieval from primary memory and secondary memory versus tasks that only require maintenance in primary memory. Specifically, when items only need to be maintained and retrieved from primary memory, processing should be phonological in nature. By contrast, when items need to be retrieved from secondary memory, processing will include additional cues (e.g., temporal-contextual cues). This suggests a parallel with transfer appropriate processing (Fisher & Craik, 1977; Morris, Bransford, & Franks, 1977), such that specific test cues will more successfully reinstate the type of processing that was engaged during one span task as opposed to another. More specific to this study, temporal-contextual cues may result in better cued recall of items processed during span tasks during which retrieval from secondary memory is necessary compared with span tasks that only require maintenance in primary memory.

Memory process approaches, including the levels of processing framework ( Craik & Lockhart, 1972; Craik & Tulving, 1975), transfer appropriate processing framework (Fisher & Craik, 1977; Morris et al., 1977), and encoding specificity principle (Thomson & Tulving, 1970; Tulving & Thomson, 1973), have regarded the type of processing, rather than the amount of rehearsal, as important for long term retention (see also Kolers & Roediger, 1984). The type of processing engaged refers to the orientation in which the participants are instructed to study the information to be recalled, such as more shallow levels of processing like phonological assessment of a word, or a more deep level of processing, such as attending to the semantic meaning of the word to remember. Multiple studies have validated the general hypothesis that a reinstatement of the original context during encoding aids in the retrieval of what was originally studied (see Roediger, Gallo, & Geraci, 2002 for a review). These studies tend to implement traditional tests of episodic memory, especially free recall, cued-recall, and recognition, and overwhelmingly support the performance advantages of reinstating the original features defined at encoding in order to aid retrieval (but see Mulligan & Lozito, 2006). Complex span tasks are, however, typically conducted over brief periods of time and include recall of unrelated items; tasks such as the operation span typically lend themselves to simple strategies, such as rote rehearsal (Dunlosky & Kane, 2007). This suggests that semantic retrieval cues in particular would not be effective in eliciting unrelated information that was only briefly presented during a complex span task. However, the transfer appropriate processing framework still has implications for simple span versus complex span tasks, and whether the respective underlying processes at encoding and retrieval are qualitatively different.

Glenberg and Swanson (1986) proposed that successful retrieval of information is due largely to the distinctiveness of the temporal context in which the information was studied. Retrieval, they argued, is aided by the cues that define a search set from which previously studied items are accessed. When that search set includes more items, the probability of recall will be lower because of the competitive nature of the selection process. In other words, if more items are

activated and included in a search set, cue overload results (Watkins & Watkins, 1975). Hence, the temporal-contextual elements of the search set should be more strictly defined in order to use those cues successfully during retrieval. Similarly, Mäntylä and Nilsson (1988) showed that the distinctiveness of the cues moderates the predicted effect of the encoding specificity principle, such that distinct, focused cues used at study, as opposed to spontaneously derived cues, are most effective at retrieval. Coupled with the literature that supports temporal-contextual processing as evidenced by errors in recall, such as incorrectly recalling items from previously presented lists (prior list intrusions; Davis, Geller, Rizzuto, & Kahana, 2008; Zaromb et al., 2006) and switching the original order when recalling in serial order (input and output transpositions; Unsworth & Engle, 2006b), this idea suggests that a memory task (e.g., a complex span task) that requires a cue-dependent search of secondary memory will also involve encoding strategies that are primarily temporal-contextual in nature.

### *Current Experiments*

The following experiments provided an application of the transfer appropriate processing framework and the temporal distinctiveness theory to understand the processes that underlie performance on simple and complex span tasks. Experiments 1 and 3 tested these ideas through the manipulation of different types of cues (i.e., temporal-contextual, phonological, or semantic) given at a delayed cued recall test of items processed in either a complex span task or a simple span task. In Experiments 2 and 4, the items processed during either a complex span task or a simple span task reappeared in a delayed memory characteristics questionnaire, which required participants to rate the how well they remembered the temporal-contextual, phonological, and semantic characteristics of each item. In other words, these experiments ascertained which kinds of characteristics are phenomenologically more memorable depending on whether the item was processed during a complex span or a simple span task. Additionally, in all four experiments, half of the items processed in both complex span and simple span tasks required immediate recall (at the standard retrieval attempt at the end of each trial), while the other half of the trials of each

task were followed by a short distracter task instead. Hence, the design for each experiment was a 2 (span task type: simple, complex) X 2 (immediate recall condition: immediate recall, no immediate recall) X 3 (delayed cue/characteristic type: temporal-contextual, phonological, semantic) within-subjects design, with percent cued recall performance or rated memorability as the relevant dependent variables.

Congruent with the transfer appropriate processing framework, retrieval of information from a complex span task should be better than a simple span task when that information is retrieved using temporal-contextual cues because the original processing during a complex span task required a search of secondary memory. In other words, a task that requires a temporally-guided search of secondary memory to maintain information during the study phase will show benefits in later performance when that initial study context is reinstated with temporal-contextual cues at retrieval. Likewise, a task that does not necessitate retrieval from secondary memory during encoding, like a simple span trial of a short list length, should show less of a benefit in performance in using temporal-contextual cues, because these trials only involve maintenance within primary memory. Similarly, temporal-contextual characteristics should be rated as more memorable for recognized items that were originally processed during a complex span task compared with a simple span task. An important application of Unsworth and Engle's primary-secondary memory hypothesis was whether performance on delayed cued recall and rated memorability of items from longer list lengths of simple span would be similar to that of a complex span task. According to the primary-secondary memory framework, longer list lengths of simple span would be more akin to complex span tasks in that the limits of primary memory would be exceeded in attempting to maintain items from longer list lengths. Hence, the following series of experiments tested these hypotheses through delayed cued recall and characteristics questionnaires of items from complex span and simple span tasks.

## Experiment 1

The research questions for Experiment 1 focus on whether simple span tasks differ from complex span tasks in the types of information encoded during study, which may be reflected in the kinds of retrieval cues that best elicit the items during a delayed cued recall test. Note that in this experiment, the simple span trials will consist of four items per trial, and thus these trials should not require retrieval from secondary memory. It is hypothesized that temporal-contextual cues will provide better access to items from a complex span task, specifically the operation span task. Immediate recall condition should not affect delayed cued recall performance of items from the operation span task because subjects should engage in retrieval from secondary memory during the encoding of those items, regardless of whether the unexpected immediate recall test occurs or not.

### *Method*

*Participants.* Sixty young adults aged 18 to 25, who were recruited from the CSU subject pool (enrolled in PSY 100 or PSY 250), participated as a course requirement. Participants were randomly assigned to one of the 12 counterbalances for this experiment, following a 3 (cue type: temporal-contextual, phonological, semantic) X 2 (immediate recall condition: recall, no recall) X 2 (span task: operation span, word span) within subjects design.

*Materials and Procedure.* An initial arithmetic task contained similar single-digit multiplication problems (e.g.,  $7 \times 4 = 28$  ?) that were to appear on the operation span task. The arithmetic task was given to participants in order to become familiar with the processing components used in the operation span, as well as to assess how long participants would take to complete the math problems that would be used as the inter- item processing element on the operation span without including the to-be-remembered (TBR) item. Participants were instructed

to read each problem out loud and verify whether the answer to the problem is “true” or “false.” The experimenter advanced the screen once a response was given. Numbers 4 to 9 were used in various combinations in order to compose the problems.

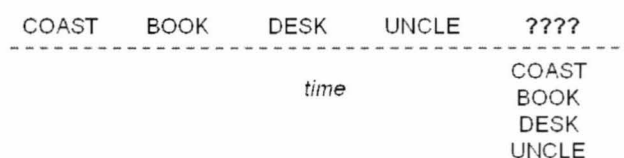
Another practice task, the matching task, served to familiarize the participants with another element that was to be included on operation span and word span trials that did not require immediate recall. In this task, participants were presented with 2-digit numbers on-screen and verified whether both numbers were even or not. If both digits were even, participants were to respond, “yes.” If both digits were odd or one was even and the other was odd, participants were to respond, “no.” The time taken to respond “yes” or “no” to each number was recorded.

Following the practice tasks, participants completed six randomly presented blocks of three word span and three operation span trials, all consisting of four TBR items each. Both the word and operation span trials were randomly presented within each test, and each test was followed by a distracter task and then a delayed cued recall test. For all of the 18 word span trials (three per test), one concrete word was presented individually at a time on the screen for 1000 ms, with participants instructed to read silently. At the end of the trials, half required immediate recall (see Figure 3a), while the other half of the trials had participants respond “yes” or “no” to five double digits numbers presented individually for 2000 ms, as they had earlier during the matching task (see Figure 3b). The operation span task also included 18 trials and was identical in procedure to the word span, except for the processing element (multiplication problem) that appeared in-between each TBR item. A multiplication problem (e.g.,  $7 \times 4 = 28$  ?) appeared on screen for 3500 ms, and participants were required to read the problem aloud and respond either “true” or “false” to the problem within the allotted time. Afterward, a TBR item appeared on screen for 1000 ms, which was read silently. Following the presentation of four problems and four words, participants were either prompted to recall the items they had seen (see Figure 4a) or to solve the matching task problems that they had practiced earlier in the experiment (see Figure 4b). Given that there were an odd number of trials of word span and operation span at each test



interval (i.e., three trials each), the trials could not be evenly split between immediate recall and no immediate recall for the tasks. In order to solve this issue, for example, two of the word span trials in three of the tests would require immediate recall while only one of the operation span trials would, whereas this pattern would reverse for the other three tests. Hence, by the end of all six tests, there was an even number of trials that required immediate recall or no immediate recall for each span task. Whether the trial required immediate recall or not was also counterbalanced.

a) **Word Span: Immediate Recall**

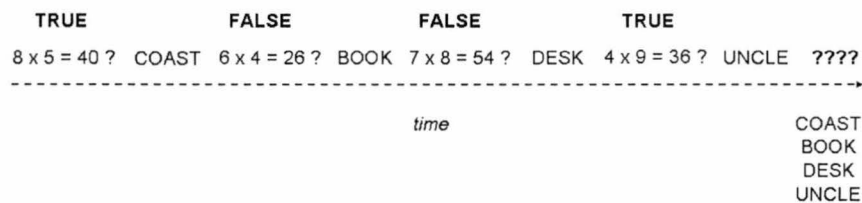


b) **Word Span: No Immediate Recall**



Figure 3. Example of (a) a word span trial with an immediate recall cue and (b) a word span trial followed by a matching task (no immediate recall).

a) **Operation Span: Immediate Recall**



b) **Operation Span: No Immediate Recall**

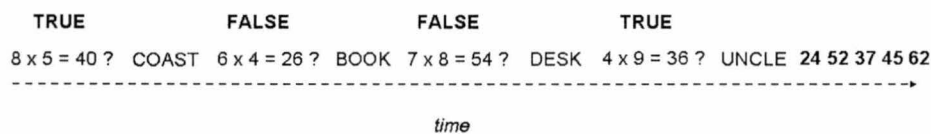


Figure 4. An example of (a) an operation span trial with an immediate recall cue and (b) an operation span trial followed by a matching task (no immediate recall).

All TBR items for this and the subsequent experiments were drawn from the Nelson, McEvoy, and Schreiber (1998) norms, which indexes forward associative strength between cue-target pairs. Forward associative strength is the likelihood that a given cue will elicit a target word response. For example, the cue “bride” prompts the recall of the target “groom” approximately 86.5% of the time that the target is cued, whereas the “acorn” cues the target “nut” 12.2% of the time. In order to create the semantic cues, targets and cues with a forward associative strength between 0.047 and 0.132 were used to avoid correct guesses ( $M = 0.079$ ,  $SD = 0.019$ ). Backwards associative strength was held constant at 0 for each cue-target pair. Cue and target frequency also ranged between 40 and 200 (cue:  $M = 92.38$ ,  $SD = 42.431$ ; target:  $M = 94.57$ ,  $SD = 43.22$ ) according to the MRC Psycholinguistic Database. In order to create phonological cues, rhyming nouns were selected, particularly those that did not share an obvious orthographic overlap (i.e., the word “teacher” cueing “creature”). Each block did not contain words that had any semantic or phonological overlap with one another, so that each cue referenced only one correct target word within that block. All words were counterbalanced for cue type, span task, and immediate recall condition.

Participants were tested at six different intervals, so that there were six trials per delayed cued recall test. Hence, 24 words total were included and viewed during each of the tests and half of those words were cued on the delayed cued recall test. Before each delayed cued recall test, the participants were given a distracter task, beginning with a demographics questionnaire after the first test, which was timed in order to administer the same amount of time for each of the other tasks serving as a delay. The other delayed tasks included word searches and listing the 50 U.S. states. After several minutes of a distracter task, words were cued using a temporal-contextual cue (e.g., “AMOUNT came before the word \_\_\_\_”), a phonological cue (e.g., “CREATURE rhymes with the word \_\_\_\_”), or a semantic cue (e.g., “AUTHORITY is related to the word \_\_\_\_”). There were four of each type of cue randomly presented at each delayed cued recall test. Items

were always cued from the second and fourth positions of each immediate recall and no immediate recall trial. These serial positions were cued because the first serial position could never be cued with a temporal-contextual cue (i.e., there is no word presented before the target to serve as a cue). The third serial position was also not cued to avoid using a target as a temporal-contextual cue for another target. For a visual layout of how each delayed cued recall test was run for each test, see Figure 5.

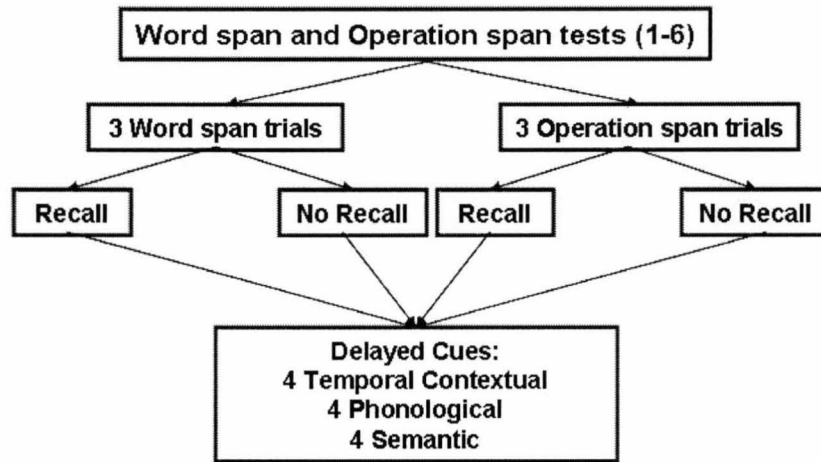


Figure 5. Design of word span and operation span tasks in Experiment 1.

### Results

See Table 1 for descriptive statistics of average response times to and errors on the processing components during operation span and the matching task that served as the no recall condition task. For future reference, all significant effects and interactions reported have met the  $p < .05$  criterion; all other  $p$  values of marginal effects will be reported specifically.

*Table 1.* Response times, errors, and time outs during the processing component and matching task

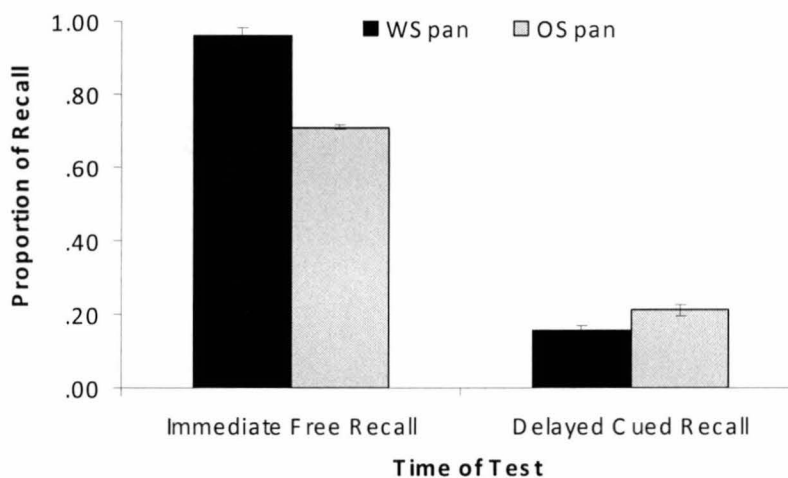
Experiment	Task	Response Time	Errors	Time Outs
1	Operation Span Processing	2875.62 (224.25)	3.67 (3.20)	10.17 (8.04)
	No Recall Matching Task	1091.04 (151.65)	1.15 (1.44)	1.77 (2.81)
2	Operation Span Processing	2988.76 (307.05)	2.88 (2.31)	20.13 (23.95)
	No Recall Matching Task	1171.22 (136.90)	2.63 (4.59)	2.67 (4.23)
3	No Recall Matching Task	1074.71 (180.84)	1.00 (1.65)	1.26 (2.92)
4	No Recall Matching Task	1066.16 (145.20)	0.92 (1.32)	0.72 (1.23)

Standard deviations are in parentheses.

In order to determine that slow response times and errors on the processing component of the operation span task did not influence results, an initial criterion analysis was conducted. Eleven participants who did not meet a 70% criterion of accuracy (slower response times than 3500 ms, or “time outs,” were also counted as inaccurate) were dropped from the analysis, but results did not change in comparison to the whole sample analysis. The only inconsistent finding was that the main effect of immediate recall condition for temporal-contextual cues became significant when including all participants, whereas it was not significant in the criterion analysis. This, however, was not theoretically relevant, and so shall not be considered further.

The first comparison was between the two times of recall of items from either span task. Note that the type of test was different between times of recall, such that the immediate test asked participants to recall in serial order (but free recall is reported) and the delayed test asked participants to retrieve items from the previous test given different cues for those items. When comparing immediate recall versus delayed cued recall performance for both span tasks, there were significant main effects of time of test,  $F(1, 59) = 1655.26$ ,  $MSE = .02$ ,  $\eta^2 = .97$ , and span task,  $F(1, 59) = 37.44$ ,  $MSE = .02$ ,  $\eta^2 = .39$ , as well as a cross-over interaction between the two variables,  $F(1, 59) = 158.35$ ,  $MSE = .01$ ,  $\eta^2 = .73$  (see Figure 6). Hence, although operation span items were initially more difficult to immediately retrieve, they were more likely to be

successfully cued than word span items at delay ( $t'(59) = -3.28, d = .52$ ), thereby replicating the delayed recall effect (McCabe, 2008).



*Figure 6.* Immediate free recall and delayed cued recall of items from word span and operation span.

Three 2-way ANOVAs were also run for each cue type on the delayed test using the whole sample. When investigating delayed cued recall performance using temporal-contextual cues, there was a significant main effect of span task such that operation span items were more likely to be recalled than word span items,  $F(1, 59) = 37.88, MSE = .01, \eta^2 = .39$ . There was also a significant main effect of recall condition, such that items that were immediately retrieved were also more likely to be cued at delay,  $F(1, 59) = 5.72, MSE = .01, \eta^2 = .09$ . Importantly, there was no interaction between the two variables,  $F < 1$ , suggesting that temporal-contextual processing was important for retrieving operation span items over word span items regardless of the immediate retrieval attempt. This result can be seen in Figure 7.

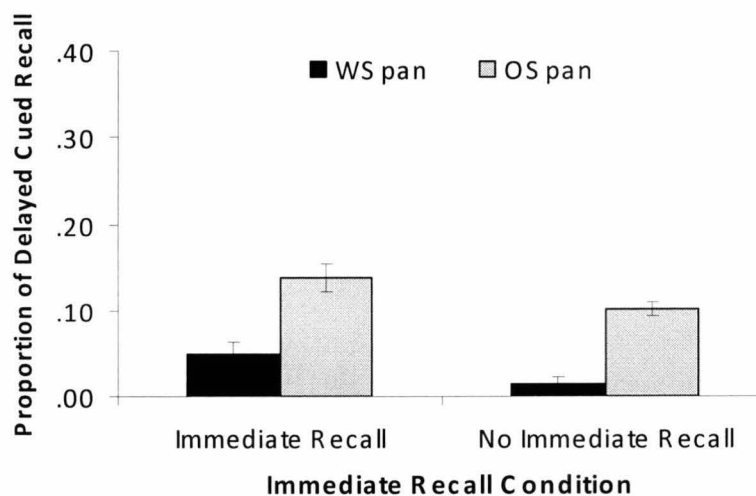


Figure 7. Delayed performance on temporal-contextual cues of operation span and word span items.

The second 2-way ANOVA for phonological cues showed a similar pattern: there were significant main effects of span task,  $F(1, 59) = 4.28$ ,  $MSE = .03$ ,  $\eta^2 = .07$ , and recall condition,  $F(1, 59) = 6.49$ ,  $MSE = .05$ ,  $\eta^2 = .10$ , and the interaction was not significant,  $F(1, 59) = 2.94$ ,  $MSE = .03$ ,  $p = .09$ ,  $\eta^2 = .05$ . Although the interaction was marginally non-significant, the difference between word span and operation span items that were not immediately recalled was statistically significant,  $t'(59) = -2.80$ ,  $d = .46$ , suggesting that immediate recall of word span items serves to boost phonological processing at encoding, lending those items to be more effectively cued with phonological cues. These results can be seen in Figure 8.

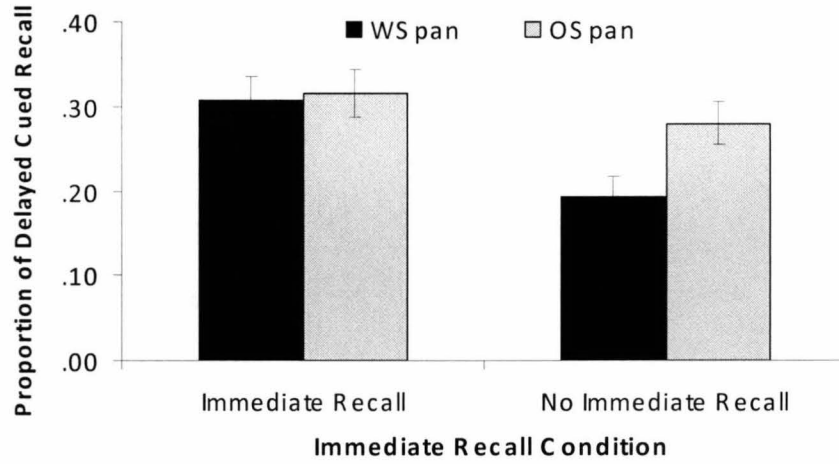


Figure 8. Delayed performance on phonological cues of word span and operation span items.

Finally, the items receiving semantic cues at delay were also submitted to a 2 x 2 ANOVA, which yielded only a significant main effect of span task,  $F(1, 59) = 16.55$ ,  $MSE = .02$ ,  $\eta^2 = .22$ . The main effect of recall condition,  $F(1, 59) = 2.26$ ,  $MSE = .02$ ,  $p = .14$ ,  $\eta^2 = .04$ , as well as the interaction,  $F < 1$ , were not significant. See Figure 9.

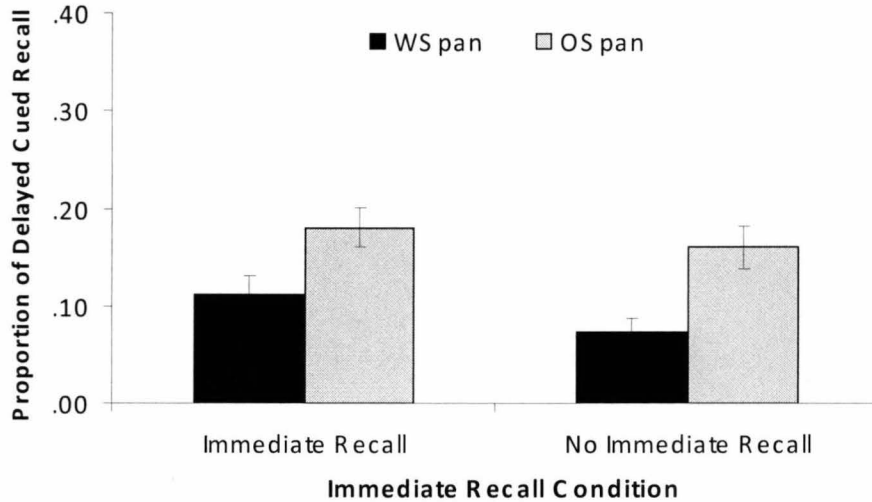


Figure 9. Delayed performance on semantic cues of word span and operation span.

Discussion

The overall results indicate that although performance suffered for immediate recall of operation span items in comparison to word span items, operation span items were more likely to be successfully cued at delay across cue types. Moreover, immediate recall condition did not interact with cued recall performance for the different span task items, suggesting that the type of processing that is important at encoding remains so regardless of the immediate retrieval attempt. This is consistent with the idea that operation span items require temporal-contextual processing when those items have been displaced from primary memory by inter-item processing tasks, whereas word span items are not effectively retrieved using temporal-contextual cues because such processing was not necessary to maintain those items only within primary memory.

There were no other specific predictions regarding delayed cued performance for the other cues, but it was interesting to note that phonological cueing of word span items matched that of operation span performance if those items had been immediately recalled. This at least partly indicates that recall from simple span tasks is more reliant on phonological cues used during immediate retrieval, and that delayed performance is hindered if the simple span items were never submitted to an immediate recall test. There was also a main effect of span task for semantic cues, which may be due to the possibility that operation span items overall were more available in long term memory. Hence, these items may be more retrievable with deeper cues, whereas word span items are better accessed with shallow cues (e.g., phonological cues).

In sum, Experiment 1 showed that objective measures of memory performance indicate that temporal-contextual cues are utilized at the encoding phase of a complex span task, and such processing can be re-instantiated to effectively retrieve the items of those tasks that require a search of secondary memory as opposed to items that remain within primary memory.

Experiment 2 attempted to provide converging evidence for this conceptual framework by investigating subjective reports regarding the types of information that are available following complex span versus simple span items.



## Experiment 2

Experiment 2 was conducted in order to investigate the subjective characteristics of the items that are re-presented at delay (Johnson, Foley, Suengas, & Raye, 1988). Are certain types of characteristics more phenomenologically memorable to a subject in one span task than the other? Furthermore, would the immediate recall condition moderate the strength of certain characteristics as memorable for items of either span task? I expected that temporal-contextual characteristics would be more memorable for operation span items. It was not expected that immediate recall condition would interact with span task or cue type, indicating that the type of processing during the encoding phases of either task is independent from the immediate recall condition.

### *Method*

*Participants.* Twenty-four participants were recruited from the CSU subject pool (students enrolled in PSY 100 and PSY 250) in exchange for course credit. Participants were young adults with normal or corrected vision. These participants had not participated in the previous experiment, and were randomly assigned to one of the four counterbalances of the experiment.

*Materials and Procedure.* The design of Experiment 2 was nearly identical to Experiment 1 (i.e., using word span and operation span with immediate recall and no immediate recall conditions). The key distinctive feature between them was the delayed characteristics questionnaires rather than the delayed cued recall tests.

Participants again were asked to complete the same arithmetic and matching tasks at the beginning of the experiment in order to familiarize themselves with the components that would be essential to the later span tasks. The word span and operation span tasks again consisted of 18

trials of four TBR items each, which were divided into three different test intervals, and each was followed by a delayed characteristics questionnaire. Within each interval, participants again were prompted to either recall the items they had seen in that trial or to solve five matching task problems at the end of each trial. After 12 trials of both recall and no recall trials (six of each span task type), participants were given a questionnaire after a distracter-filled delay. All 48 of the words reappeared one at a time, which were each followed by a three-question memory characteristics questionnaire regarding the item's place on the list (i.e., "How much can you remember about WHERE IT WAS ON THE STUDY LIST (e.g., the word that came before or after it, whether it was at the beginning or end of a trial)?"), its physical characteristics (i.e., "How much can you remember about the PHYSICAL CHARACTERISTICS (i.e., what the word sounded like, what the word looked like)?"), and the meaning of the word (i.e., "How much can you remember about the MEANING (e.g., the definition of the word, your emotional reaction to the word, related words that the word made you think of)?"). The order of the presentation of the questions was random, and subjects responded to each question on a 1 to 9 Likert scale. See Figure 10 for a visual layout of the proposed design of Experiment 2.

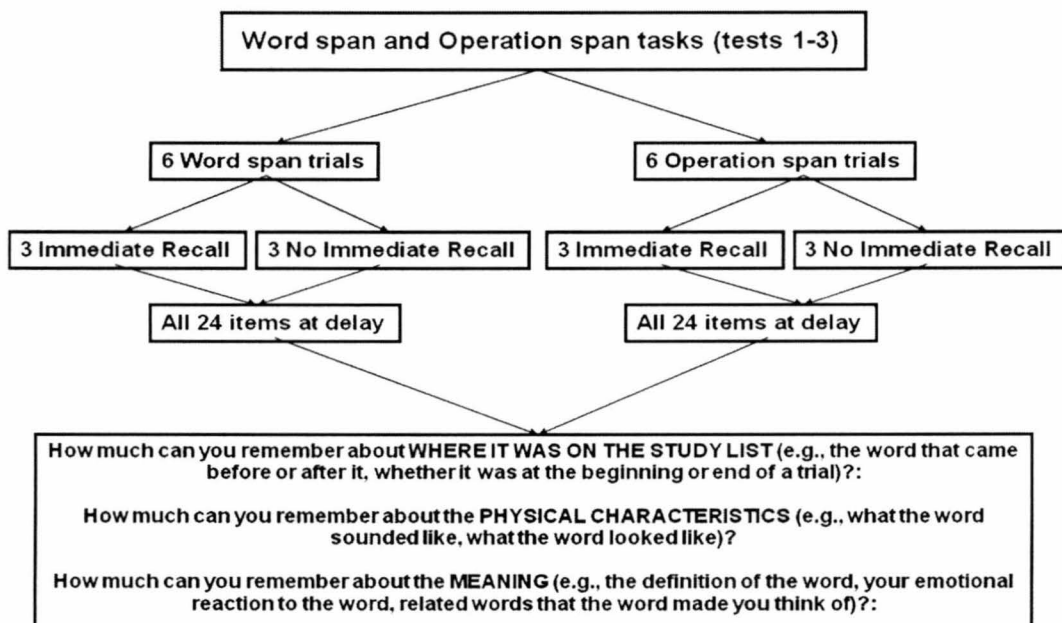


Figure 10. Design of word span and operation span tasks in Experiment 2.

## Results

Means and standard deviations for the processing component and matching task response times can be found in Table 1.

In order to determine that the results of the whole-sample analysis were not due to individual subjects' errors and slowed response times during the operation span task's processing component, an analysis was run with only those subjects who met a 70% criterion for accurate responses before time had run out on the task. This led to dropping five subjects from this criterion analysis. When these data were excluded, the same pattern of results emerged that was found using the whole sample, and so all data points were included for the following analysis.

The data were once again submitted to three 2 (span task: operation span, word span) x 2 (recall condition: immediate recall, no immediate recall) ANOVAs for each of the three characteristic types. For ratings on temporal-contextual characteristics, there were significant main effects of both span task,  $F(1, 23) = 46.52$ ,  $MSE = .40$ ,  $\eta^2 = .67$ , and immediate recall condition,  $F(1, 23) = 33.87$ ,  $MSE = .16$ ,  $\eta^2 = .57$ , showing that operation span items and immediately recalled items were more likely to be rated as memorable. However, these two variables did not interact,  $F < 1$ . Similarly, phonological characteristics' memorability ratings showed a similar pattern of results, such that the main effects of span task,  $F(1, 23) = 20.37$ ,  $MSE = .33$ ,  $\eta^2 = .47$ , and immediate recall condition,  $F(1, 23) = 10.50$ ,  $MSE = .23$ ,  $\eta^2 = .31$ , were significant in the same directions; however, the interaction was not reliable,  $F < 1$ . Finally, the semantic characteristics also showed the same patterns, such that operation span items were rated as more semantically memorable than word span items,  $F(1, 23) = 37.11$ ,  $MSE = .32$ ,  $\eta^2 = .62$ , and that immediately recalled items were rated as more semantically memorable,  $F(1, 23) = 3.20$ ,  $MSE = .19$ ,  $\eta^2 = .41$ . Once again, these variables did not interact,  $F < 1$ . Because recall condition never interacted with the span task variable, the data were collapsed across recall condition in

Figure 11. Note that although ratings were always higher for operation span items, the effect size was largest for temporal-contextual characteristics.

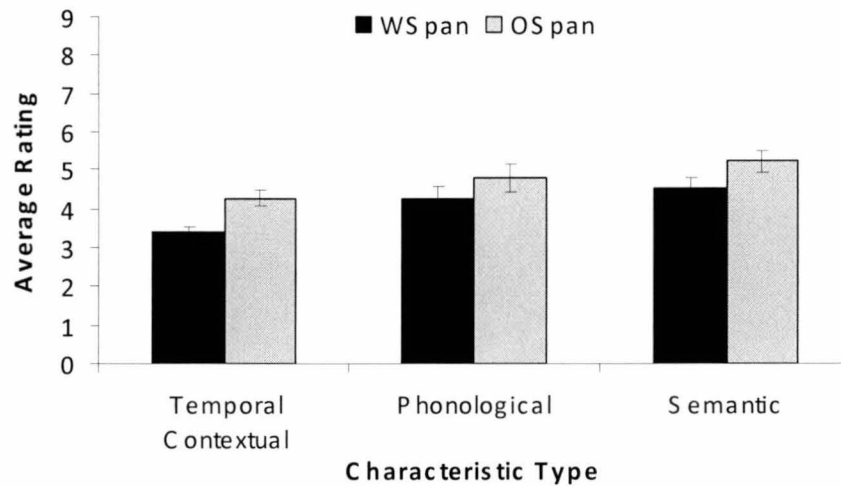


Figure 11. Average ratings given for temporal-contextual, phonological, and semantic characteristics of items from word span and operation span.

### Discussion

The results of Experiment 2 complement the results of Experiment 1, indicating that not only are temporal-contextual cues important for retrieval of operation span items, but also that participants actually report remembering more about operation span items than word span items, especially temporal-contextual characteristics. Because word span items are presumably still in primary memory when using list lengths up to four items, participants would not need to engage in any kind of refreshing or shifting process to maintain those items during the encoding phase. Hence, temporal-contextual processing would not be as enhanced as it would be for successful maintenance of items of a complex span task. Given transfer appropriate processing, the temporal-contextual search in order to shift items back into primary memory during the encoding phase of the operation span task becomes important for both cued recall of those items as well as what people report they can remember about those items. More importantly, the lack of interaction with immediate recall condition showed that temporal-contextual characteristics were

important for operation span regardless of whether there was a retrieval attempt, and further illustrates the significance of the actual refreshing act during the task in which participants must engage.

Now that the first two experiments have shown evidence for more temporal-contextual processing during operation span as compared to word span, an important extrapolation of the primary-secondary memory hypothesis is that longer list lengths of a simple span task are more similar to complex span tasks in terms of the kind of processing that is utilized to accomplish the task. Hence, temporal-contextual processing should also be important for simple span items that have been displaced to secondary memory due to new incoming items within a trial. Experiment 3 tested this very hypothesis.

### Experiment 3

Experiment 3 replicated the design of Experiment 1 with the exception that it included only the word span at sub- and supra-span list lengths. This experiment tested the idea proposed by Unsworth and Engle (2006a; 2007b) that simple span tasks at longer list lengths are more akin to complex span tasks in that they require retrieval from secondary memory. Hence, I expected that there would be an overall greater proportion of items recalled from longer list lengths at a delay, although the opposite would be true at immediate recall. I also expected to find the same pattern of results from Experiment 1, such that temporal-contextual cues would better access word span items at longer list lengths. This finding would be consistent with the idea that items that are displaced from primary memory are subjected to more temporal-contextual processing during the study phase in order for successful recall.

#### *Method*

*Participants.* The study recruited 54 participants from the CSU psychology subject pool (students enrolled in PSY 100 and PSY 250). All participants were young adults with normal or corrected vision, and had not participated in the previous two experiments. Participants were randomly assigned to one of 18 counterbalances (three per counterbalance).

*Materials and Procedure.* The materials and procedure in Experiment 3 were similar to those in Experiments 1 and 2, except that only word span trials of list lengths 4 and 8 were used. Participants completed the same matching task that was used in Experiments 1 and 2 in order to prepare for trials that did not require immediate recall during the word span task. Because the operation span task was not used in Experiment 3, no arithmetic task was necessary to practice the multiplication problems.

The word span task consisted of 24 trials, half of which presented four TBR items and the

other half of those trials presented eight TBR items. The same words from previous experiments served as stimuli and were presented on screen for 1000 ms each in both list length conditions. Similarly to the first two experiments, half of the trials of both list lengths prompted participants to recall the words that were immediately presented, while the other half of the trials of both list lengths required no immediate recall and instead were followed by a matching task sequence of five double-digit numbers presented one at a time for 2000 ms each. The presentation of both types of trials was randomized, and the task in its entirety was separated into six different delayed cued recall tests, which themselves were randomly presented. Two trials from list length 4 (one immediate recall trial and one no immediate recall trial) and two trials from list length 8 (one immediate recall trial and one no immediate recall trial) were included in each of the six delayed cued recall tests. Of the 24 words viewed during each test interval, 12 total items were cued after a distracter-filled delay. Again, items were cued using the same temporal-contextual, phonological, and semantic cues that were used for Experiment 1. Items from list length 4 were always cued from the second and fourth serial positions, and items from list length 8 were always cued from the second, fourth, sixth, and eighth serial positions for the same reasons stipulated in Experiment 1. Figure 12 shows the design for each of the six different tests.

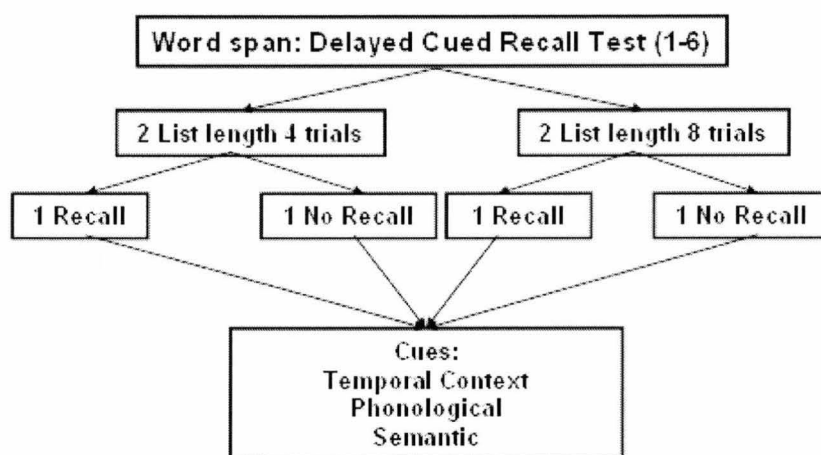


Figure 12. Design of delayed cued recall tests of word span items in Experiment 3.

## Results

Due to the fact that there were no processing components involved in Experiment 3, it was not necessary to ensure that the obtained data were not due to participants' errors and response times. See Table 1 for errors and response times for the matching task, which was still used for the no immediate recall trials during the span task.

*List Length 4 and List Length 8 Analyses.* First, an analysis comparing performance for immediate free recall and delayed cued recall was performed to investigate whether the delayed recall effect would be found for list length 8 items. There was a main effect of time of test,  $F(1, 53) = 1897.74$ ,  $MSE = .01$ ,  $\eta^2 = .97$ . There was also a main effect of list length, with more items recalled from list length 4 overall,  $F(1, 53) = 446.26$ ,  $MSE = .01$ ,  $\eta^2 = .89$ . Finally, there was a significant interaction between time of test and list length,  $F(1, 53) = 454.93$ ,  $MSE = .01$ ,  $\eta^2 = .90$ , but this was not the cross-over interaction that is indicative of the delayed recall effect found in McCabe (2008). As can be seen in Figure 13, delayed cued recall performance was statistically equivalent between list lengths.

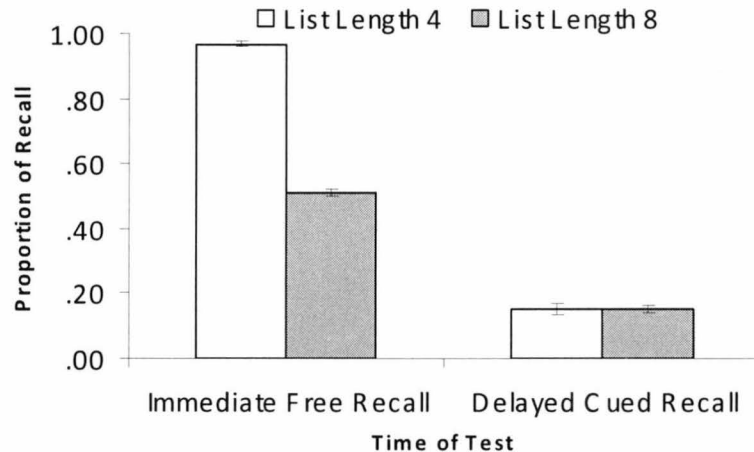


Figure 13. Immediate recall and delayed cued recall of items from list length 4 and list length 8 trials of word span.

Performance given each of the cue types was analyzed using 2 (list length: 4, 8) x 2 (immediate recall condition: immediate recall, no immediate recall) repeated measures ANOVAs. When analyzing temporal-contextual cues, there was a significant main effect of list length,  $F(1,$



53) = 4.63,  $MSE = .01$ ,  $\eta^2 = .08$ , but in the opposite direction than predicted; specifically, items from list length 4 trials were more likely to be recalled than list length 8 items. There was no main effect of immediate recall condition nor was there an interaction,  $F_s < 1$ . When the analysis was conducted for performance on the phonological cues, the main effects of list length and immediate recall condition, as well as the interaction, were not significant,  $F_s < 1$ . Finally, only a marginally significant effect of immediate recall condition resulted when analyzing semantic cues,  $F(1, 53) = 2.90$ ,  $MSE = .02$ ,  $p = .09$ ,  $\eta^2 = .05$ . The main effect of list length and the interaction were not significant,  $F_s < 1$ . See Figures 14 for these comparisons.

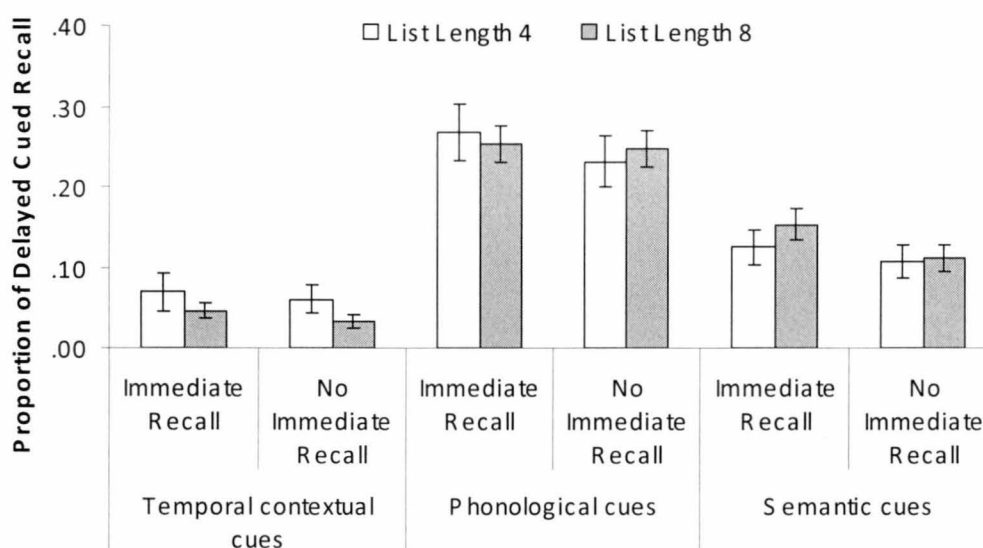


Figure 14. Delayed cued recall performance across recall condition and cue type for items from list lengths 4 and 8.

*List Length 4 and the First Four Items of List Length 8 Analyses.* The following analyses were executed under the hypothesis that when primary memory exceeds four items, the earlier presented items will be displaced into secondary memory. Perhaps comparing list length 4 items to list length 8 items is not a fair comparison because the later presented items in list length 8 trials would theoretically still remain in primary memory at recall. Hence, it is pertinent to

compare the delayed cued recall of the first four items of list length 8 to list length 4 items using the same analyses.

First, the potential replication of McCabe's (2008) delayed recall effect was investigated. There was a main effect of list length, where once again there was better performance overall for list length 4 trials than for the first four items of list length 8,  $F(1, 53) = 54.70$ ,  $MSE = .02$ ,  $\eta^2 = .51$ . There was also a main effect of time of test,  $F(1, 53) = 1659.60$ ,  $MSE = .02$ ,  $\eta^2 = .97$ . Finally, there was an interaction between list length and time of test,  $F(1, 53) = 83.72$ ,  $MSE = .01$ ,  $\eta^2 = .61$ . Although there was a numerical advantage for the first four items of list length 8 at delayed cued recall, this difference was not significant,  $t'(53) = .78$ , *ns*. See Figure 15 for these data. It is noteworthy, however, that there was a marked improvement in immediate recall performance when comparing list length 8 items in total to the first four items of those same trials,  $t'(53) = -7.85$ ,  $d = 1.38$ .

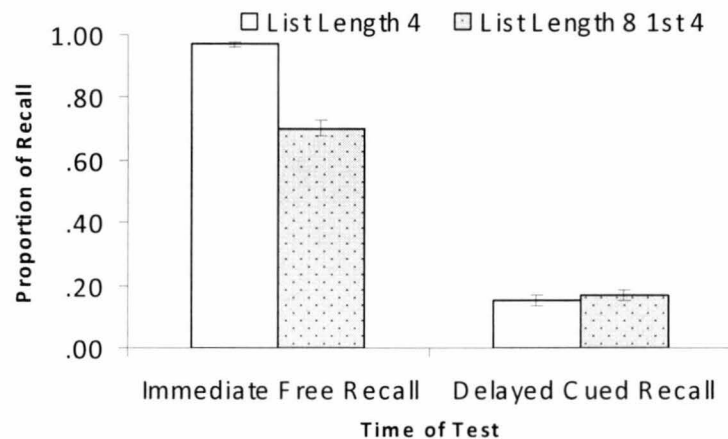


Figure 15. Immediate recall and delayed cued recall of list length 4 items and the first four items of list length 8 of word span.

Once again, the delayed cued recall data were submitted to three 2 x 2 ANOVAs for each of the three cue types. For temporal-contextual cues, both main effects of list length and immediate recall condition, as well as the interaction between the two variables, were not significant, all  $F_s < 1$ . This was also consistent for phonological cues,  $F_s < 1$ . Finally, semantic

cued performance showed only a marginal effect of recall condition,  $F(1, 53) = 2.66$ ,  $MSE = .02$ ,  $p = .10$ ,  $\eta^2 = .05$ , but the effect of list length and the interaction were not significant,  $F_s < 1$ . See Figure 16 for a graphic display of these comparisons.

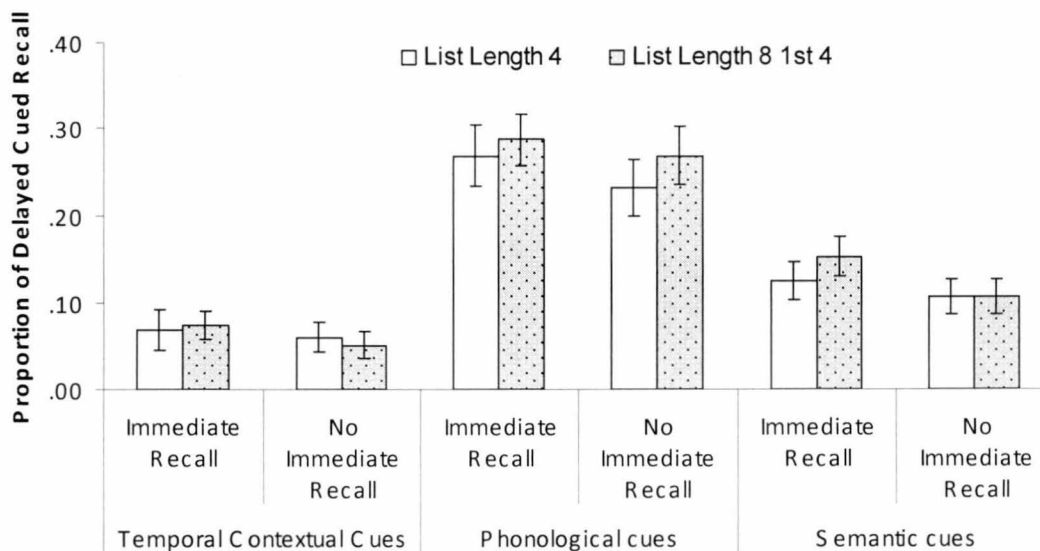


Figure 16. Delayed cued recall across recall condition and cue type for list length 4 items and the first four items of list length 8.

### Discussion

First, the delayed recall effect (McCabe, 2008) was not replicated in either analysis of the overall recall performance of the list lengths. Furthermore, the only relevant difference found in the previous analyses was that there was a main effect of list length for temporal-contextual cues, yet this was in the opposite direction than what was predicted. Specifically, items in list length 4 trials benefited more from temporal-contextual cues than list length 8 items. Although this seems inconsistent with the previous experiments and the primary-secondary memory hypothesis, an even more interesting finding is that this effect was not found when delayed cued recall performance was considered only for the first four items of those list length 8 trials. Hence, it could be argued that the null effect of list length in the first analysis may have occurred because of a larger fan associated with the increased number of words in the trial as opposed to sub-span

list lengths of word span, rather than because those items became more temporally distinctive from a search of secondary memory.

The comparison of the first four items of list length eight was an important analysis in particular given the primary-secondary memory hypothesis that list lengths exceeding four items will begin to shift first-presented information into secondary memory in order to successfully maintain those items. Hence, those items in particular should have benefited more from temporal-contextual cues if they had actually been refreshed into primary memory through a temporal-contextual search of secondary memory during encoding. Due to the null effect, it is not clear if this process did not occur at all or was simply less evident through the given temporal-contextual cues. Nevertheless, it still poses a problem for the primary-secondary memory framework in positing that longer list lengths of simple span tasks are similar to complex span tasks. In order to determine whether this null effect was an artifact of the type of delayed measure used, it was pertinent to investigate whether participants would find items from longer list lengths of simple span more phenomenologically memorable than items from shorter list lengths; this was done in Experiment 4.

## Experiment 4

Experiment 4 attempted to replicate the results of Experiment 2 with the sub-and supra-span list lengths of word span described in Experiment 3. Again, I investigated whether certain types of characteristics of an item are more phenomenologically memorable to a person as a function of shorter versus longer list lengths of word span. Based on the primary-secondary memory hypothesis, I expected that temporal-contextual characteristics should be rated as more memorable for recognized items of longer list lengths of word span. This result was not expected to interact with immediate recall condition, under the assumption that a retrieval attempt is not necessary to engage in a secondary memory search during the actual encoding, which is what would yield in higher ratings for temporal-contextual characteristics.

### *Method*

*Participants.* Thirty-six participants (six per counterbalance) were recruited from the CSU psychology subject pool (enrolled in PSY 100 and PSY 250) to participate in the experiment in exchange for partial course credit. All participants were young adults aged 18 to 25 with normal or corrected to normal vision, and no one had participated in the previous experiments.

*Materials and Procedure.* Experiment 4 was similar to the previous experiments, especially to Experiment 3, except that there were three delayed characteristics questionnaires for all of the items of both list lengths 4 and 8 of the word span task. Again, all participants began with the practice matching task to familiarize themselves with the task that was to be used during the no immediate recall trials during the word span task.

The word span was similar to Experiment 3, where trials of four and eight TBR items were randomly presented, and half of the trials of each list length were followed by either a recall attempt or a sequence of five matching task problems, presented for 2000 ms each. The word

span task itself was segmented into three different study periods that included four trials of list length 4 (two immediate recall and two no immediate recall trials) and four trials of list length 8 (two immediate recall and two no immediate recall trials). After a distracter-filled delay following each study interval, a delayed characteristics questionnaire was given, which included all 48 of the items from the preceding span task. The re-presentation of items was followed by a three-question memory characteristics questionnaire, which was identical to the one given in Experiment 2, and inquired about the temporal-contextual, phonological, or semantic characteristics remembered on a scale of 1 to 9 for each individual item. The presentation of the questions was random. Figure 17 presents a visual layout of the design of Experiment 4 for both list lengths of the word span task.

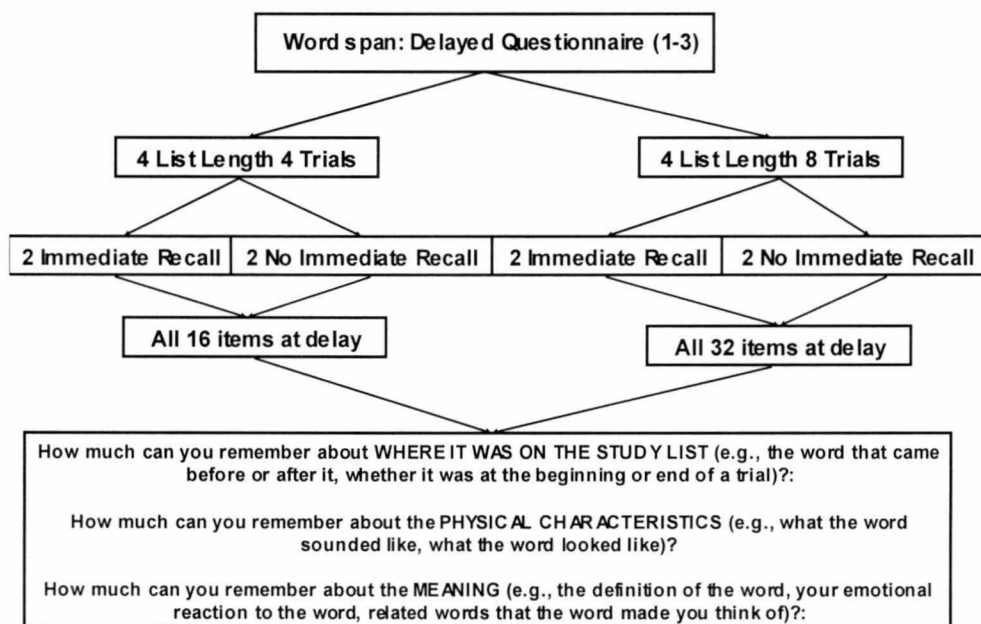


Figure 17. Design of word span trials in Experiment 4.

### Results

The average response times, errors, and number of time outs for the matching task are included in Table 1. All participants were included in the following analyses. Consistent with

Experiment 3, results are divided between ratings on both list lengths as well as the first four items of list length 8.

*List Length 4 and List Length 8 Analyses.* The ratings for the stimuli were submitted to a 2 (list length: 4, 8) X 2 (recall condition: immediate recall, no immediate recall) repeated measures ANOVA for each characteristic type. For temporal-contextual characteristics, participants rated previously recalled items as more memorable than those that were not submitted to an immediate recall attempt,  $F(1, 35) = 21.75$ ,  $MSE = .24$ ,  $\eta^2 = .38$ . However, the main effect of list length and the interaction were not significant,  $F_s < 1$ . Participants also found phonological characteristics to be more memorable for immediately recalled items,  $F(1, 35) = 10.92$ ,  $MSE = .09$ ,  $\eta^2 = .24$ , as well as items from list length 4 as opposed to list length 8,  $F(1, 35) = 4.95$ ,  $MSE = .12$ ,  $\eta^2 = .12$ . However, the interaction was not significant,  $F(1, 35) = 1.39$ , *ns*. Finally, semantic characteristics, the effect of immediate recall condition was once again significant,  $F(1, 35) = 25.73$ ,  $MSE = .10$ ,  $\eta^2 = .42$ , but the main effect of list length,  $F(1, 35) = .41$ , *ns*, and the interaction,  $F(1, 35) = 1.74$ , *ns*, were not significant. See Figure 18.

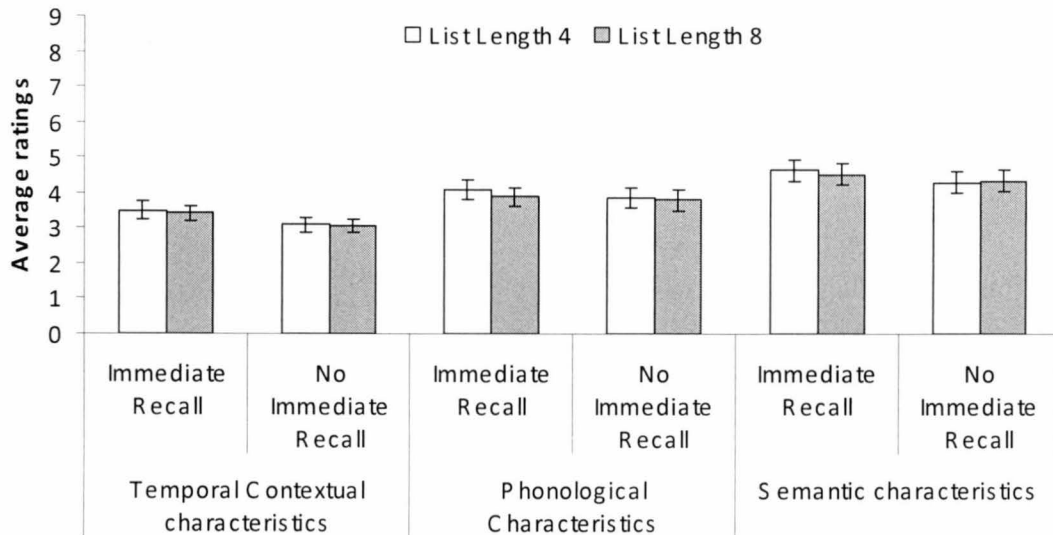


Figure 18. Average ratings on the delayed characteristics questionnaire across recall condition and cue type for items from list length 4 and list length 8.

*List Length 4 and First Four Items of List Length 8 Analyses.* The same ANOVAs were run but using the first four items of list length 8 in comparison to the list length 4 items. For temporal-contextual characteristics, there was once again a main effect of recall condition,  $F(1, 35) = 14.57$ ,  $MSE = .32$ ,  $\eta^2 = .29$ , yet the main effect of list length and interaction were not significant,  $F_s < 1$ . The main effect of immediate recall condition was also significant when comparing ratings on phonological characteristics,  $F(1, 35) = 4.99$ ,  $MSE = .12$ ,  $\eta^2 = .13$ . However, the main effect of list length,  $F(1, 35) = 1.67$ , *ns* and the interaction,  $F(1, 35) = 2.50$ , *ns* were not significant. Finally, the semantic ratings made at delay showed a significant effect of immediate recall condition,  $F(1, 35) = 14.54$ ,  $MSE = .12$ ,  $\eta^2 = .29$ , and no significant effect of list length,  $F < 1$ . However, there was a marginally significant interaction between the two variables,  $F(1, 35) = 3.82$ ,  $MSE = .14$ ,  $p = .06$ ,  $\eta^2 = .10$ . See Figure 19 for these results.

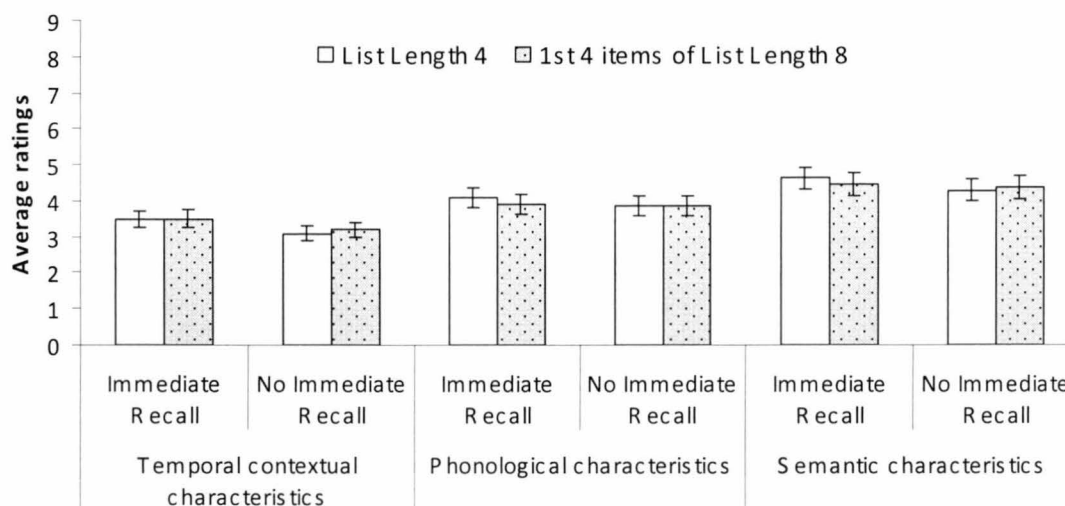


Figure 19. Average ratings on the delayed characteristics questionnaire across recall condition and cue type of items from list length 4 and the first four items of list length 8.

### Discussion

The significant results of the experiment showed that people tend to find items that were recalled at the immediate test as most memorable for each characteristic type. This effect was



significant in both analyses comparing list length 4 and list length 8 as well as the comparison between list length 4 and the first four items of list length 8. Importantly, these effects did not interact with list length, suggesting that participants did not find items more memorable if they were in trials that required displacement into secondary memory. The only marginally significant difference between list length types was for phonological ratings in the first set of analyses comparing list length 4 and list length 8 items, but there were no specific theoretical expectations pertaining to this particular finding. Furthermore, this effect was absent in the second analyses utilizing only ratings on the first four items of list length 8. There was a marginally non-significant interaction between list length and immediate recall condition for semantic ratings, but this again was neither expected nor relevant. Hence, no support was found for the expectation that participants would find temporal-contextual characteristics more memorable when they concern items that they conceivably may have displaced and retrieved from secondary memory. This finding, however, at least complements the objective delayed cued recall from Experiment 3, wherein no differences were found between list lengths.

## General Discussion

The present series of experiments suggests that temporal-contextual processing is important for retrieval of items that were initially encoded during a complex span task, but not a simple span task, even at longer list lengths of the simple span. This was evident through both objective measures of delayed cued recall (Experiment 1) as well as the subjective characteristics that people report remembering about those items during operation span (Experiment 2).

Although operation span items were more likely to be successfully cued and more phenomenologically memorable across characteristic type, the effect of span task was generally larger for temporal-contextual cues and characteristics than any others ( $\eta^2 = .39$  and  $\eta^2 = .67$  for Experiments 1 and 2, respectively). Moreover, recall condition did not change performance on temporal-contextual cueing for the different span tasks, nor was it important for memorability of temporal-contextual characteristics of operation span items, suggesting that a retrieval attempt is not responsible for differences in temporal-contextual processing for operation span compared to word span, but that the process is more specific to the encoding of those items.

These measures were employed again in Experiments 3 and 4, respectively, for word span of longer list lengths compared to shorter list lengths, in order to determine whether temporal-contextual processing is also important for items that are hypothetically displaced from primary memory by more recently presented items. If simple span tasks at longer list lengths encourage the same type of processing as complex span tasks in utilizing temporal-contextual processing to retrieve displaced information in secondary memory, then the same benefit of temporal-contextual cues should have been found for list length 8 items as it had been found for operation span items. There were no differences, however, for objective cued recall or subjective memorability ratings, even when comparing the first four items of list length 8 trials against list

length 4 items. This general finding has a number of implications for the primary-secondary memory hypothesis, temporal distinctiveness models of retrieval, and working memory research in general.

#### *Primary-Secondary Memory Hypothesis*

As stated, there was no evidence in the last two experiments for the idea that simple span at longer list lengths would lead to processing that was similar to complex span tasks. Although there is very clear support of Unsworth and Engle's (2006b) idea that the complex span tasks require temporal-contextual processing during an encoding phase, this finding was not extended to longer list lengths of simple span.

Support for the primary-secondary memory framework comes from findings that simple span performance for list lengths greater than four both load on a different factor than simple span at lower list lengths and also predicts a unique amount of variance on higher order cognition, like general fluid intelligence (Unsworth & Engle, 2006a). We (McCabe, Loaiza, & Roediger, 2009) have replicated this finding in our own lab, and others have identified that secondary memory measures are uniquely important for the relationship between working memory capacity and fluid intelligence (De Alwis et al., 2009; Mogle, Lovett, Stawski, & Sliwinski, 2008). The act of retrieving information from secondary memory appears to be the reason that tasks tax working memory capacity, and cause them to be predictive of other measures. This follows closely with Cowan's (1993) hypothesis that the focus of attention can only maintain approximately four chunks of information. Hence, simple span tasks appear similar to complex span tasks: the act of displacement during longer simple span trials, when new incoming information exceeds four items, better predicts other higher cognitive abilities much as complex span tasks do. It appears that these longer trials of simple span, too, could tax working memory capacity. The presented series of experiments was important for testing this assumption, not by investigating whether predictive utility would be similar between span task types, but that the actual *processes* elicited by the tasks would also be similar. This was evidently not the case.

If supra-span list lengths of simple span are both separable from sub-span list lengths and can predict a similar unique amount of variance in variables like fluid intelligence, then why wouldn't the processes during these trials and complex span trials be similar? Which features make the span task types both similar enough to predict variance in other variables? Why wouldn't that shared feature be the degree of temporal-contextual processing during encoding? While the limitations of the current series of experiments only allow for conjecture, there are testable answers to these questions that are further informed by literature on this topic.

*Similarities and Differences Between Types of Span Tasks.* First, although sub- and supra-span list lengths of simple span loaded on different factors, a peculiar finding is that the supra-span list length simple span trials did not load on the same factor as complex span trials (Unsworth & Engle, 2006a). Furthermore, a significant proportion of the variance was unique to supra-span list lengths and complex span trials, suggesting that while both types of trials are significantly predictive of fluid intelligence, there is still a large degree of variance that is not shared between them. Hence, these two findings suggest that these factors are still separable both in their loadings and the way they predict fluid intelligence, which may be due to the different types of processing engaged by the tasks.

McCabe (2008) also found that the greater degree of covert retrieval processing during complex span tasks in the act of switching between completing the processing component and then searching secondary memory to maintain displaced items during the task serves to boost long-term memory performance. A very basic indication of this idea is the cross-over interaction between the time at which people are tested on items and the type of span task they completed. This interaction was found in Experiment 1, but was not replicated in Experiment 3, which suggests that covert retrieval processes may not have been engaged in the same way as they are for complex span tasks. Ongoing work in our lab has indicated that this covert retrieval process may more specifically be identified as attentional refreshing, which may facilitate temporal-

contextual processing with the increasing number of opportunities to refresh TBR items (McCabe & Loaiza, 2009).

*Attentional Refreshing and Evidence for Temporal-Contextual Processing.* A number of recent studies (e.g., Camos, Lagner, & Barrouillet, 2009; Hudjetz & Oberauer, 2007; Saito & Miyake, 2004) have identified that attentional refreshing, or re-activating just-presented information, is imperative for performance on working memory tasks and is qualitatively different than the more domain-specific act of articulatory rehearsal. In our own lab, the act of attentional refreshing appears to maintain information by facilitating temporal-contextual processing, such that earlier temporal positions of items receive more refreshing opportunities than later positions, resulting in an upward linear trend between refreshing opportunities and delayed recall performance. Hence, refreshing and temporal-contextual processing appear to be intrinsically tied to one another.

It is not clear given the present data that the act of refreshing and therefore enhancement of temporal-contextual information occurs for supra-span list length trials of simple span, which may be an important reason for why temporal-contextual cues were unimportant for delayed cued performance. Unsworth and Engle (2006b) also found evidence of temporal-contextual processing in complex span given findings of increased temporal-contextual errors in complex span recall for low spans as opposed to high span individuals. Unsworth and Engle (2008) also identified focus switching to be an important factor in the relationship between working memory and general fluid intelligence. Countless studies have also investigated recall performance of items across serial position given different factors of immediate and delayed free recall, in continuous distracter paradigms, and analyses of different types of errors.

One important recent study by Unsworth, Heitz and Parks (2009) investigated recall of consonant trigrams following different retention intervals, and specifically manipulated the amount of time preceding specific trials to identify the degree to which temporal distinctiveness is important for recall. The study found that a greater amount of time in-between trials lead to a

rebound in recall performance in comparison with a control condition that showed the typical rate of forgetting with increasing delay before retrieval. This indicates that temporal distinctiveness of items from other TBR items facilitates retrieval, and generally supports Glenberg and Swanson's (1986) proposal that temporally distinct contexts help to constrain searches and avoid temporal confusability.

An implication for the current results suggests that there may have been greater temporal confusability of simple span items, regardless list length, because those trials did not afford the same benefit of temporal definition from inter-item processing elements that occurred for complex span items. For example, Unsworth and Engle (2007b) have shown that a greater number of transposition errors, or switching the original order of presented items during recall, are made during simple span tasks as opposed to complex span tasks. However, a number of serial position studies have found evidence of temporal associations within recall of items from simple word lists. For example, Kahana and colleagues (e.g., Howard & Kahana, 2002; Kahana, Howard, Zaromb, & Wingfield, 2002; Zaromb et al., 2006) have repeatedly noted a lag recency effect in immediate free recall, such that recall of a word from a list will cue words from nearby serial positions. Hence, it might not be that there are no temporal associations made in standard list learning scenarios, but that temporal-contextual processing is better facilitated at encoding of complex span items, and thus more accessible to temporal-contextual retrieval cues due to temporally distinguishing inter-item processing elements that require people to continually refresh information in working memory.

### *Limitations*

There were several limitations which could be addressed in future work. The temporal-contextual cues used in the experiments also are susceptible to criticism. Temporal processing is sometimes considered dissociable from ordinal processing, or the order in which items are presented (Farrell & McLaughlin, 2007). It could be said that the temporal-contextual cues would qualify more as ordinal cues, in the sense that I used serial position information in order to cue

items in nearby serial positions. A response to this criticism is that temporal distinctiveness still appeared to be an important factor in the series of experiments in that the temporal-contextual cues used only worked for operation span items, whereas the same cues used for supra-span list lengths of word span may not have been as temporally defined. Hence, if the cues used were more ordinal in the sense that they were cueing item-based information with serial position information, temporal distinctiveness still ended up being a more important factor that was accessed.

Temporal-contextual cues at other levels of the hierarchy may have been important as well, and perhaps could have better elicited any potential temporal-contextual processing that had been executed during supra-span list lengths of word span. For example, some data from our lab have suggested that people access temporal information more at a trial level than an item-based level. Unsworth and Engle (2007a) have also argued that there are different hierarchies of temporal-contextual processing, such as at the list context or the word context. This may be a worthwhile line of research in specifying at which level of hierarchical temporal-contextual processing is important during a complex span task.

Tulving's (1983) idea of "retrieval mode" suggested that people must adopt a mental set in order to engage in recovering memory traces. This may extend to a plural form, such that there can be multiple retrieval modes that are dependent on a type of processing, especially given the idea of transfer-appropriate processing/encoding specificity principle. Specifically, when in a temporal-contextual "retrieval mode" that is instantiated with a cue at delay, such a mental set could be hindered by switching modes, or cues, to other access other types of information, like phonological or semantic information. Hence, it may be more pertinent to separate cue types between tests rather than presenting them randomly within a test.

#### *Future Directions*

A number of experiments could be executed to address some of the unresolved issues of this study. Given the present series of experiments, we can only conjecture on the potential

interactions between complex span and supra-span list lengths of simple span tasks. It would be pertinent to include both types of trials within the same experiment to test whether the benefit of delayed cued performance persists for complex span items even when comparing more directly to supra-span list lengths of simple span. Future studies could also investigate trials where word span items are temporally isolated with just unfilled delays after each word to determine whether temporally isolating those items without event-filled delays after each item (as in operation span or other complex span tasks) will yield in the same degree of temporal distinctiveness.

These results should also be replicated with other types of complex and simple span tasks to verify that temporal-contextual processing is also important for all span tasks that are purported to tax working memory. It was clear in Experiments 1 and 2 that operation span items are more accessible and memorable than word span items across cue and characteristic type, which suggests that searching secondary memory does not only utilize temporal-contextual cues, but other cues may also benefit retrieval as well. For example, a reading span task may utilize semantic cues in addition to temporal-contextual cues in a similar way because depending on administration of the task, the TBR items may be given extra semantic processing due to their nature of being in the sentence serving as a processing component. However, if temporal-contextual processes are the defining factor of working memory, it should still be evident that all complex span tasks share these processes. All other processes that are particular to the task would be considered as domain-specific processing.

The presented findings also have many implications for research of individual differences in working memory. Unsworth and Engle (2006b) have already identified that temporal-contextual processing during complex span tasks is diminished in low span individuals, such that they are more likely to make errors that are indicative of temporal confusability. Furthermore, individual differences in focus switching also are important for the relationship between working memory and fluid intelligence (Unsworth & Engle, 2008). It can be inferred that low span individuals may not constrain their search sets as effectively as high span individuals. This may



lead to less effective use of temporal-contextual cues at delay given that temporal associations during encoding are weaker in low span individuals. Furthermore, this difference in temporal-contextual processing may be important for the relationship between working memory capacity and other higher order cognitive abilities like general fluid intelligence. Hence, diminished ability to engage in temporally constrained searches of secondary memory could be added as a possible explanation of span differences across individuals, along with other explanations like deficits in inhibition of irrelevant information (Lustig, May, & Hasher, 2001) and diminished executive attention (Engle & Kane, 2004).

Similarly, there are also implications for research investigating declines in working memory performance with aging. The ability to refresh information within the focus of attention has been identified as important deficiency in older adults in comparison to young adults (Johnson, Reeder, Raye, & Mitchell, 2002; Verhaeghen & Hoyer, 2007). Additionally, Kahana et al. (2002) found that older adults are not able to maintain the temporal associations between freely recalled items, evidenced by shallower lag recency curves. In addition, older adults have an increased rate of prior list intrusions, showing that they do not have the same temporally defined search set that younger adults utilize during recall. Thus, there is evidence that both focus switching and temporal-contextual processing are deficient in older adults, and future investigations could determine whether these two deficiencies are closely related to one another.

### *Conclusions*

It is clear that complex span tasks elicit temporal-contextual processing during the encoding phase of the task, and given transfer appropriate processing, it is important to re-instantiate those processes in order to access previously studied information at delay. The argument that simple span tasks also encourage temporal-contextual processing similar to complex span tasks was not supported from either objective or subjective indices. It is unclear whether temporal-contextual cues and characteristics are more important for complex span tasks because of displacement into secondary memory, or if they are important because items studied in

complex span tasks are more temporally distinguishable and allow for a greater degree of temporal association than simple span trials of any length. Still, the present study is an important step in identifying differences between span tasks through experimental manipulation of cues and phenomenological characteristics.

## References

- Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation*, (pp. 47-89). New York: Academic Press.
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*, 457-469.
- Cowan, N. (1993). Activation, attention, and short term memory. *Memory and Cognition*, *21*(2), 162-167.
- Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, *9*, 143-148.
- Craik, F. I. M., Gardiner, J. M., & Watkins, M. J. (1970). Further evidence of the negative recency effect in free recall. *Journal of Verbal Learning and Verbal Behavior*, *9*, 554-560.
- 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.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268-294.
- Davis, O. C., Geller, A. S., Rizzuto, D. S., & Kahana, M. J. (2008). Temporal associative processes revealed by intrusions in paired-associate recall. *Psychonomic Bulletin and Review*, *15*(1), 64-69.
- De Alwis, D., Myerson, J., Hershey, T., & Hale, S. (2009) Children's higher-order cognitive abilities and the development of secondary memory. *Psychonomic Bulletin and Review*, *16*(5), 925-930.
- Delaney, P. F., & Sahakyan, L. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory and Cognition*, *35*(5), 1074-1082.
- Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227-1245.
- Ebbinghaus, H. (1885/1913). *Memory*. (H. A. Ruger & C. E. Bussenius, Transl.) New York: Teachers College, Columbia University.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (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*(3), 309-331.
- Farrell, S., & McLaughlin, K. (2007). Short-term recognition memory for serial order. *Memory and Cognition*, *35*(7), 1724-1734.
- Fisher, & Craik, F. I. M. (1977). Interaction between encoding and retrieval operations in cued recall. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(6), 707-711.
- Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 3-15.
- Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal-context. *Journal of Mathematical Psychology*, *46*, 269-299.
- Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span task paradigm. *Memory and Cognition*, *35*(7), 1675-1684.
- Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*(4), 371-376.
- Johnson, M. K., Reeder, J. A., Raye, C. L., & Michell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychological Science*, *13*(1), 64-67.
- Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(3), 530-540.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent variable approach to verbal and visuo-spatial memory span and reasoning. *Psychonomic Bulletin and Review*, *9*, 637-671.
- Kolers, P., & Roediger, H. (1984). Procedures of mind. *Journal of Verbal Learning & Verbal Behavior*, *23*(4), 425-449.
- Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, *130*(2), 199-207.
- Mäntylä, T., & Nilsson, L.-G. (1988). Cue distinctiveness and forgetting: Effectiveness of self-generated retrieval cues in delayed recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 502-509.

- McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, *58*(2), 480-494.
- McCabe, D. P., & Loaiza, V. Factors dissociating retrieval from working memory and episodic memory: Dual effects of covert retrieval. Manuscript in preparation.
- McCabe, D. P., Loaiza, V., & Roediger, H. L. Explaining age differences in fluid intelligence using estimates of primary and secondary memory derived from working memory span tasks. Manuscript under review.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519-533.
- Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological Science*, *19*(11), 1071-1077.
- Mulligan, N. W., & Lozito, J. P. (2006). An asymmetry between memory encoding and retrieval: Revelation, generation, and transfer-appropriate processing. *Psychological Science*, *17*(1), 7-11.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation/>.
- Roediger, H. L., Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, *10*(5/6), 319-332.
- Saito, S., & Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in reading span performance. *Journal of Memory and Language*, *50*, 425-443.
- Shulman, H. G. (1971). Similarity effects in short-term memory. *Psychological Bulletin*, *75*(6), 399-415.
- Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, *86*(2), 255-262.
- Tulving, E. (1983). Ecphoric processes in episodic memory. *Phil. Trans. R. Soc. Lond. B., Biological Science*, *302*(1110), 361-370.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352-373.
- Unsworth, N., & Engle, R. W. (2006a). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*, 68-80.
- Unsworth, N., & Engle, R. W. (2006b). A temporal-contextual retrieval account of complex span: An analysis of errors. *Journal of Memory and Language*, *54*, 346-362.
- Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory

capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114(1), 104-132.

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

Unsworth, N., & Engle, R. W. (2008). Speed and accuracy of accessing information in working memory: An individual differences investigation of focus switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), 616-630.

Unsworth, N., Heitz, R. P., & Parks, N. A. (2009). The importance of temporal distinctiveness for forgetting over the short term. *Psychological Science*, 19(11), 1078-1081.

Verhaeghen, P., & Hoyer, W. J. (2007). Aging, focus switching, and task switching in a continuous calculation task: Evidence toward a new working memory control process. *Aging, Neuropsychology, and Cognition*, 14, 22-39.

Watkins, M. J. (1974). Concept and measurement of primary memory. *Psychological Bulletin*, 81(10), 695-711.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, 104, 442-452.

Wundt, W. (1897/1969a). *Outlines of Psychology*. (C. H. Judd, Transl.) St. Clair Shores, MI: Scholarly Press.

Wundt, W. (1912/1973). *An Introduction to Psychology*. New York: Arno Press.

Zaromb, F. M., Howard, M. W., Dolan, E. D., Yevgeniy, S. B., Tully, M., Wingfield, A., & Kahana, M. J. (2006). Temporal associations and prior-list intrusions in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 792-804.