INDIVIDUAL DIFFERENCES IN COMPLEX MEMORY SPAN AND EPISODIC RETRIEVAL: EXAMINING THE DYNAMICS OF DELAYED AND CONTINUOUS DISTRACTOR FREE RECALL

A Dissertation
Presented to
The Academic Faculty

by

David I. Unsworth

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Experimental Psychology

Georgia Institute of Technology

May 2006

Individual Differences in Complex Memory Span and Episodic Retrieval:
Examining the Dynamics of Delayed and Continuous Distractor Free Recall

Approved by:

Dr. Randall W. Engle, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Daniel Spieler
School of Psychology
Georgia Institute of Technology

Dr. Paul Corballis
School of Psychology
Georgia Institute of Technology

Dr. Anderson D. Smith
School of Psychology
Georgia Institute of Technology

Dr. David Washburn
Department of Psychology
Georgia State University

Date Approved: December 11, 2005

ACKNOWLEDGEMENTS

I would like to acknowledge my committee members, Randy Engle, Dan Spieler, Paul Corballis, Andy Smith, and David Washburn for their insightful suggestions and helpful comments. I would also like to thank Rich Heitz and Tom Redick for helpful discussions and encouragement in all phases of this project. Finally, I would like to thank my family for all of their encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Individual differences on complex memory spans predict a variety of higher-order cognitive tasks (e.g. reading comprehension, reasoning, following direction) as well as low-level attention tasks (e.g. Stroop, dichotic listening, antisaccade). The current study attempted to better determine the role of individual differences in complex memory span and episodic retrieval. Specifically, two experiments explored the possibility that individual differences in complex memory span reflect differences in the ability to successfully retrieve items from secondary memory via a cue-dependent search process. High and low complex span participants were tested in delayed (Experiment 1) and continuous distractor (Experiment 2) free recall with varying list-lengths. Across both experiments low spans recalled fewer items than high spans, recalled more previous list intrusions than high spans, and recalled at a slower rate than high spans. It is argued that low spans search through a larger set of items than high spans and, thus low spans' episodic retrieval deficits are associated with an inability to use cues to guide a search and retrieval process of secondary memory. Implications for dual-component models of memory are discussed.

CHAPTER 1: INTRODUCTION

Complex memory span tasks such as reading (Daneman & Carpenter, 1980) and operation span (Turner & Engle, 1989) have been shown to be important predictors of a number of higher-order and lower-order cognitive processes. In these tasks to-be-remembered items are interspersed with some form of distracting activity such as reading sentences or solving math operations. In terms of higher-order cognitive processes, these complex spans have been shown to predict reading comprehension (Daneman & Carpenter, 1980), fluid reasoning (Ackerman, Beier, & Boyle, 2002; Conway et al., 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004; Kyllonen & Christal, 1990; Unsworth & Engle, 2005a), vocabulary learning (Daneman & Green, 1986), performance on the Scholastic Aptitude Test (Turner & Engle, 1989), and intentional learning (Kyllonen & Stephens, 1990; Unsworth & Engle, 2005b). In terms of lower-order cognitive processes, these complex span tasks have been shown to predict performance on attention and inhibition tasks like Stroop (Kane & Engle, 2003; Long & Prat, 2002), dichotic listening (Conway, Cowan, & Bunting, 2001), antisaccade (Kane, Bleckley, Conway, & Engle, 2001; Unsworth, Schrock, & Engle, 2004) and flankers (Heitz & Engle, 2005; Redick & Engle, in press). Furthermore, these tasks have been shown to predict important phenomena such as early onset Alzheimer's (Rosen et al. 2002), life-event stress (Klein & Boals, 2001), and stereotype threat (Schmader & Johns, 2003; see Unsworth, Heitz, & Engle, 2005 for a review).

A number of theories have postulated a central mechanism as the main underlying construct responsible for the predictive power of these tasks. These include the inhibition view of Hasher, Zacks, and colleagues (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999), the controlled (or executive) attention view espoused by Conway, Engle, Kane, and colleagues (Engle & Kane, 2004; Kane, Conway, Hambrick, & Engle, in press), the capacity of attention view supported by Cowan (2001; 2005), and the more traditional view of resource sharing currently supported by Barrouillet and Camos (2001; Barrouillet, Bernardin, & Camos, 2004) among others.

The present work explored a specific possibility of individual differences in complex spans. Because complex spans are fundamentally memory tasks, the present work explored the possibility that the primary process tapped by these tasks is one of retrieval. Individual differences in complex span, therefore, are differences in the ability to effectively retrieve items. Specifically, it is argued that variation in complex span is due to differences in the ability to use cues to guide a search and retrieval process of secondary memory.

Individual Differences in Complex Span and Episodic Retrieval

Over the last few years, a number of studies have convincingly demonstrated that variation in complex span is related to variation in the ability to retrieve information from secondary memory under conditions of interference. In particular, these studies have shown that individuals who differ in complex span performance also differ in the ability to effectively retrieve information from secondary memory where competition between items is high. Under conditions of reduced competition or interference, however, complex span differences either do not appear or are greatly reduced. For instance, consider a study by Conway and Engle (1994), which examined individual differences in complex span and retrieval in a version of the Sternberg probe recognition task

(Sternberg, 1966). In this study, Conway and Engle manipulated the amount of interference present by allowing some items to repeat across sets. For example, in some conditions a target in one set of items could also be a target in another set of items, thus increasing interference across trials. In non-interference conditions, however, target items could not repeat across sets. Conway and Engle found that high and low spans did not differ in conditions of reduced interference (i.e., a new set of items on each trial), but that low spans were substantially slower in conditions of interference (i.e., when items were repeated across trials). The authors suggested that individual differences in complex span reflected differences in the ability to combat interference at retrieval.

Additional studies that have examined the relation between complex memory span and retrieval under conditions of interference have suggested similar results. For instance, using a variant of the Brown-Peterson task, Kane and Engle (2000) found that low spans showed a greater buildup of proactive interference (PI) across trials than high spans. Furthermore, Rosen and Engle (1998) found that low spans made more first-list intrusions on second-list learning in a paired-associates task than high spans. These results suggest that those participants who score high on measures of complex span tend to do better on memory retrieval measures than participants who score low on complex span measures particularly under conditions of interference. Specifically, these results suggest that low complex span individuals are more likely to recall fewer target items, recall more intruding items, and are slower to recall items than high complex span individuals.

Other studies have examined individual differences in the complex span tasks themselves and have suggested that differences in retrieval processes may be an important contributor to the predictive power of these tasks. For instance, Unsworth and Engle (in press) recently proposed a model of verbal complex and simple memory span tasks suggesting that one reason these tasks "work" is because they require retrieval from both primary and secondary memory. Similar to Cowan (2001) and Davelaar et al., (2005), Unsworth and Engle suggested that primary memory has an upper bound of approximately four items. When more items are present, items that have been displaced from primary memory must be retrieved via a search process (e.g., Shiffrin, 1970) of secondary memory.

Unsworth and Engle suggested that the reason that complex span tasks typically show consistently moderate correlations with higher-order cognition is because they require retrieval from both primary and secondary memory at all list-lengths. This is because the processing component in these tasks (e.g., solving math operations) typically displaces items from primary memory and thus requires that the items be retrieved via a search of secondary memory in the face of PI from prior trials (e.g., May, Hasher, & Kane, 1999). Unsworth and Engle showed that complex memory spans showed constant correlations with a composite of fluid abilities across all list-lengths. Unsworth and Engle (in press) took this as evidence that one important process involved in complex memory span tasks and their relation to higher-order cognition was the ability to effectively retrieve items from secondary memory. Furthermore, Unsworth and Engle argued that a possible reason for this was because, as the task progressed, low spans were unable to delimit the search set only to items from the current trial and had to search through a much larger set of items than high spans. Unsworth and Engle argued that low complex span participants are more susceptible to PI via poor list-discrimination processes than participants with higher span scores.

Additional support for this notion came from a detailed analysis of errors in the operation and reading span tasks (Unsworth & Engle, 2005c). These results suggested that individual differences in working memory capacity (WMC), as measured by the complex span tasks, were partially due to differences in retrieval from secondary memory. Specifically, the results suggested that nearly all participants correctly recalled the last item presented (although not always in the correct serial position). However, the further back in time an item was presented, the less likely it was to be recalled. Additionally, examining this by high and low scorers on the tasks suggested that low span individuals were much less likely to recall items presented further back in time and were more likely to intrude items from previous lists. Unsworth and Engle (2005c) argued for a temporal discrimination interpretation of the results as proposed by Glenberg and Swanson (1986) and Brown, Preece, and Hulme (2000). Specifically, Unsworth and Engle argued that participants were using temporally defined search sets and that low span individuals were poorer at using temporal cues to delimit the search set to only the current items. This resulted in more items being subsumed under a given retrieval cue (i.e., cue-overload, Watkins, 1979), some of which were items from previous lists. Thus, low spans had to search through larger search sets which contained some extra-list items and this resulted in a lower probability of recall for these individuals.

Based on this evidence, Unsworth and Engle (in press; 2005d) argued for a dual component model combining a flexible attentional component (primary memory) with a cue-dependent search mechanism of secondary memory. Unsworth and Engle argued that individual differences in WMC (as measured by complex span tasks) result from differences in the ability to maintain items in primary memory and/or differences in the ability to use cues to guide the search process of secondary memory. Furthermore, as noted previously, differences in susceptibility to PI arise primarily due to differences in list-discrimination processes in which low spans are unable to focus the search of secondary memory only on current target items and thus, they must search through a larger set of items than high spans. The aim of the present paper was to better explore this possibility by examining individual differences in complex span and the dynamics of free recall.

Dynamics of Free Recall

Free recall is especially suited for testing aspects of this model. Specifically, free recall studies have shown that the recency effect is limited to approximately the last four items (e.g. Watkins, 1974), that PI only affects prerecency items and not recency items (Craik & Birtwistle, 1971; Davelaar et al., 2005; Unsworth & Engle, 2005d), that measurements of recall latency and inter-response times increase as the secondary memory search set increases (Rohrer & Wixted, 1994), and inter-response times for recency items are very rapid (i.e., under one second), but inter-response times for prerecency items are much longer (i.e., over one second; Murdock, 1972). These results suggest that primary memory is limited to approximately four items, and that items within primary memory are impervious to PI (see also Cowan, Johnson, & Saults, 2005; Halford et al., 1988) resulting in fast and perfect recall of these items. Items that have been displaced from primary memory, however, must be retrieved via a search of secondary memory. This search process is hindered by a number of factors including PI, which increases the size of the search set and leads to both slower and less accurate recall of items from secondary memory.

In the present paper, the hypothesis that individual differences in complex memory span are partially due to differences in a cue-dependent search process will be examined via the random search model (Bousfield, Sedgewick, & Cohen, 1954; Kaplan, Carvellas, & Metlaly, 1969; McGill, 1963; Rohrer & Wixted, 1994; Wixted & Rohrer, 1994). In this model a retrieval cue delimits a search set that includes representations of target items as well as extraneous items. Item representations are randomly sampled from the search set at a constant rate, one item at a time (serial search). The retrieval process includes a sampling-with-replacement process such that after an item representation has been sampled and recalled the same representation still has an equal chance of being selected on the next sample. Target items that have been previously recalled, intruding items, or target items that are not recoverable, are not recalled but still can be sampled from the search set. As the retrieval process proceeds, the probability of recalling a new target item decreases because each sample is likely to be an already recalled target item or an extraneous item.

Assuming a constant sampling time per item, McGill (1963) demonstrated how this simple random sampling-with-replacement model predicted exponentially declining rates of recall and conversely predicted cumulative exponential recall curves (see also Rohrer & Wixted, 1994; Vorberg & Ulrich, 1987). Indeed, beginning with the work of Bousfield and colleagues (Bousfield & Sedgewick, 1944; see also Indow & Togano, 1970; Roediger, Stellon, & Tulving, 1977), research has found that cumulative latency distributions are well described by the cumulative exponential,

$$F(t) = N(1 - e^{-\lambda t}), \quad (1)$$

where F(t) represents the cumulative number of items recalled by time t, N represents asymptotic recall, and λ represents the rate of approach to asymptote. Using the random search model and the parameter estimates obtained from fitting the cumulative exponential to the cumulative latency distributions, several studies have shown that N and λ change as a function of different task manipulations (see Wixted & Rohrer, 1994 for a review). For instance, Herrmann and Chaffin (1976; see also Metlay, Handley, & Kaplan, 1971) showed that categorical recall from large categories resulted in more items being recalled (e.g., a larger N) and that the rate of recalling items from the larger category was slower (e.g., a smaller λ) compared to recall from small categories. This result makes perfect sense for the random search model. The search set for the large category was larger than the small category search set which resulted in more items being recalled in the large search set. At the same time, because there are more items in the large category search set, the time to sample a new item after the recall of several items should be slower than for the small search set. Indeed, Wixted and Rohrer (1993) have noted that "in a sampling-with-replacement serial search model, the average time required to find target items in a search set increases linearly with the size of that set" (p. 1036). That is, it takes more time in a large search set to find a new item that has not been recalled previously. This also implies that mean inter-response time (IRT) associated with larger search sets should be larger than mean IRT for small search sets.

Additional evidence in support of this view comes from recent studies examining latency distributions in free recall. For instance, Wixted and Rohrer (1993) had participants perform a variant of the Brown-Peterson task where the first three trials were all from the same category to see how the buildup of PI would affect the latency distributions. The authors found that as PI accrued, estimates of both N and λ decreased, suggesting that the search set increased for subsequent trials using the same category. That is, the size of the search set increased because the number of previous list representations within it increased. However, in the release from PI condition, estimates of N and λ increased slightly. These results suggested that as PI accrued, the search set became progressively larger because the search set was delimited to all category instances based on the retrieval cue. Under release conditions, the retrieval cue specified only the new category instances and thus the search set excluded items from the previous trials.

Additional work by these authors on delayed free recall tasks has demonstrated the utility of the random search model in examining the temporal aspects of free recall. For instance, Rohrer and Wixted (1994) found that increases in list-length resulted in decreases in λ and corresponding increases in mean recall latency , suggesting that as list-length increased the size of the search set increased (i.e., cue-overload). Furthermore, Rohrer and Wixted (1994) found that increasing presentation duration and presumably increasing the amount of attention paid to items at recall resulted in an increase in probability correct but no change in λ or mean recall latency . The authors suggested that this was because the presentation duration manipulation increased the likelihood that a target would be recoverable during the recall phase but left the search set unaffected. That is, most search models assume that items that have an absolute strength greater than some value can be recovered, but that items whose absolute strength falls below that value cannot be recovered. Increasing presentation duration and attention at encoding increases items' absolute strength but does not affect the size of the search set (see also Shiffrin, 1970). These and other results suggest that the random search model is a useful tool in interpreting recall performance under a variety of conditions including the effects of PI (Wixted & Rohrer, 1993), manipulations of list-length (Rohrer & Wixted, 1994), manipulations of presentation duration (Rohrer & Wixted, 1994), episodic vs. semantic recall (Rohrer, 2002), as well as a number of categorical recall findings (see Wixted and Rohrer, 1994 for a comprehensive review).

Individual Differences in Complex Span and the Dynamics of Free Recall

As an initial test of the notion that variation in complex span is due to differences in the ability to retrieve items from secondary memory via a cue-dependent search process, Unsworth and Engle (2005d) had high and low span participants perform an immediate free recall task. Participants were given 15 lists of twelve words each. Each word was presented alone for 1 s and participants were given 30 s to recall as many words as possible during the recall period. During the recall period the experimenter pressed a key each time the participant recalled a word (both target words and intrusions).

According to the framework presented previously, if high spans are better at delimiting the search set to only the current items, whereas lows spans have trouble delimiting the search set to only the current trials (list-discrimination) then high spans should recall more words than lows and their rate of approach to asymptotic recall levels should be faster than low spans. This is precisely what was found. Fitting the cumulative exponential for each individual resulted in larger N and λ estimates for high spans than for low spans.

Furthermore, examining mean recall latency and mean IRTs for each span group for the first six responses suggested that high spans had shorter mean recall latencies (4.19 s vs. 6.14 s) and IRTs (1.64 s vs. 2.65 s) than low spans. Taken together, the results suggested that low span individuals were less efficient in delimiting their search sets to only the current trials and thus had a much larger search set to search through than did high span individuals.

In addition, because the task was immediate free recall, it is possible that high and low spans differ in their ability to hold items within primary memory (c.f., Cowan, 2001). Examining serial position effects for both high and low spans suggested that highs and lows differed slightly in the recency portion of the curve (where probability correct was quite high) but that high spans had a higher probability of correctly recalling items from the prerecency portion of the list. Unsworth and Engle (2005d) also examined IRT differences between high and low spans for those items thought to be retrieved from the either primary or secondary memory. The first two IRTs, which represent items predominantly from the recency portion of the serial position curve, were very fast and roughly equivalent (e.g., IRT1 = 902 ms vs. IRT2 = 1015 ms). And this was true for both high (IRT1 = 784ms vs. IRT2 = 849ms) and low spans (IRT1 = 1020 ms vs. IRT2 = 1180 ms). This suggested that these items were recalled from primary memory for both span groups. However, beginning with the third IRT, the IRTs increased and were consistent with a random search model. Here, complex span differences began to appear with low spans demonstrating longer IRTs than high spans, suggesting that the low spans were taking more time to sample new items than high spans from secondary memory.

Rationale for the Present Study

The Unsworth and Engle (2005d) immediate free recall findings provide initial support for the notion that part of low spans' recall deficits are due to an inability to correctly delimit the search set of secondary memory compared to high spans. However, these findings are limited by the fact that some of the items were recalled (theoretically) from primary memory and, thus the results do not clearly demonstrate differences in the search process between the two groups. This is most readily seen when examining recency and prerecency differences between the two groups. Small differences between the span groups occurred for recency items in terms of both probability correct and the recall latency measures. However, large differences occurred between the groups when examining prerecency items in terms of both probability correct and recall latency. Theoretically, the search process only occurs for the prerecency items (where the largest span differences occurred) and, thus in order to get a cleaner picture of complex span differences in the search process, the role of primary memory must be minimized. In order to better examine retrieval processes in the absence of retrieval from primary memory, the dynamics of free recall in both delayed (Experiment 1) and continuous distractor (Experiment 2) free recall tasks was examined in the present experiments. Because the distractor task during the retention interval in both delayed and continuous distractor recall is thought to displace items from primary memory (e.g., Bjork & Whitten, 1974; Glanzer & Cunitz, 1966), the results should demonstrate that the recency effect is greatly reduced and that individual differences in complex span occur at all serial positions, rather than only prerecency positions.

The present set of experiments also examined how manipulations of list-length would affect the correlation between complex span and free recall. If low span individuals are partially deficient in their ability to recall due to an inability to constrain the search set, by manipulating the size of the search set we should be able to reduce the correlation between λ (rate of approach to asymptote) and performance on complex memory spans by manipulating list-length. Rohrer and Wixted (1994) showed that manipulations of list-length resulted in changes in λ. Specifically, as list-length increased, λ decreased. Because the Unsworth and Engle (2005d) immediate free recall experiment used list-lengths of twelve items, it is possible that low span individuals had trouble constraining their search sets to such a large number of items. Thus, there may be a point at which low spans' ability to constrain the search set becomes ineffective and if given fewer items they may perform equivalently to high spans. That is, the correlation between λ and performance on complex memory spans may be moderated by list-length with the largest correlation occurring for the largest list-length.

List-length was also manipulated in order to try and gauge how different high and low spans are in terms of the size of their search sets. Specifically, the manipulation of list-length should provide a means to determine where high and low spans show similar performance on the latency measures. For instance, if low spans have larger search sets than high spans the list-length manipulation can provide a rough estimate of when highs and lows have equivalent search sets. That is, we can ask "Do low spans search through the same number of items at list-length six as highs do at list-length nine?" Although the manipulation of list-length probably will not provide a precise estimate of differences in search set size, it should provide a fairly gross measure of differences.

Although the current work explores the possibility that high and low spans differences are due to differences in search set size (see below), other viable alternatives also exist. Therefore, in both Experiments four possibilities for differences between high and low span participants in retrieval based on the random search model were tested (see Figure 1). The first possibility (panel A in Figure 1) is that low spans search through a larger search set than high spans resulting in fewer target items recalled, smaller values of λ, and shorter recall latencies and faster IRT values. Based on the list-discrimination hypothesis and previous work, this possibility is considered to be the most tenable. This hypothesis suggests that low spans are less efficient than high spans at using retrieval cues to delimit the search set and thus must search through a larger search set than high spans. Because the search set contains both targets and nontargets, low spans will take more time to sample and recover new target items from the search set than high spans resulting in longer mean recall latency and slower IRTs. This is reminiscent of Wixted and Rohrer's (1993) PI finding. Additional support for this position should come from an analysis of recall errors. If low spans are poorer at list-discrimination than high spans, then low spans should recall more previous list intrusions than high spans and these intrusions should come from the immediately preceding list (e.g., Unsworth and Engle, 2005d).

The second possible (see panel B in Figure 1) reason why low span participants recall fewer items than high span participants is that low spans search a much smaller search set than high spans. Because there are fewer target items within the search set, low spans will subsequently recall fewer items and have lower values of N. However, this position suggests that low spans should actually have larger values of λ than high spans and consequently should have shorter mean recall latency and faster mean IRT values than high spans. Additionally, this scenario seems possible given that previous research examining the temporal dynamics of free

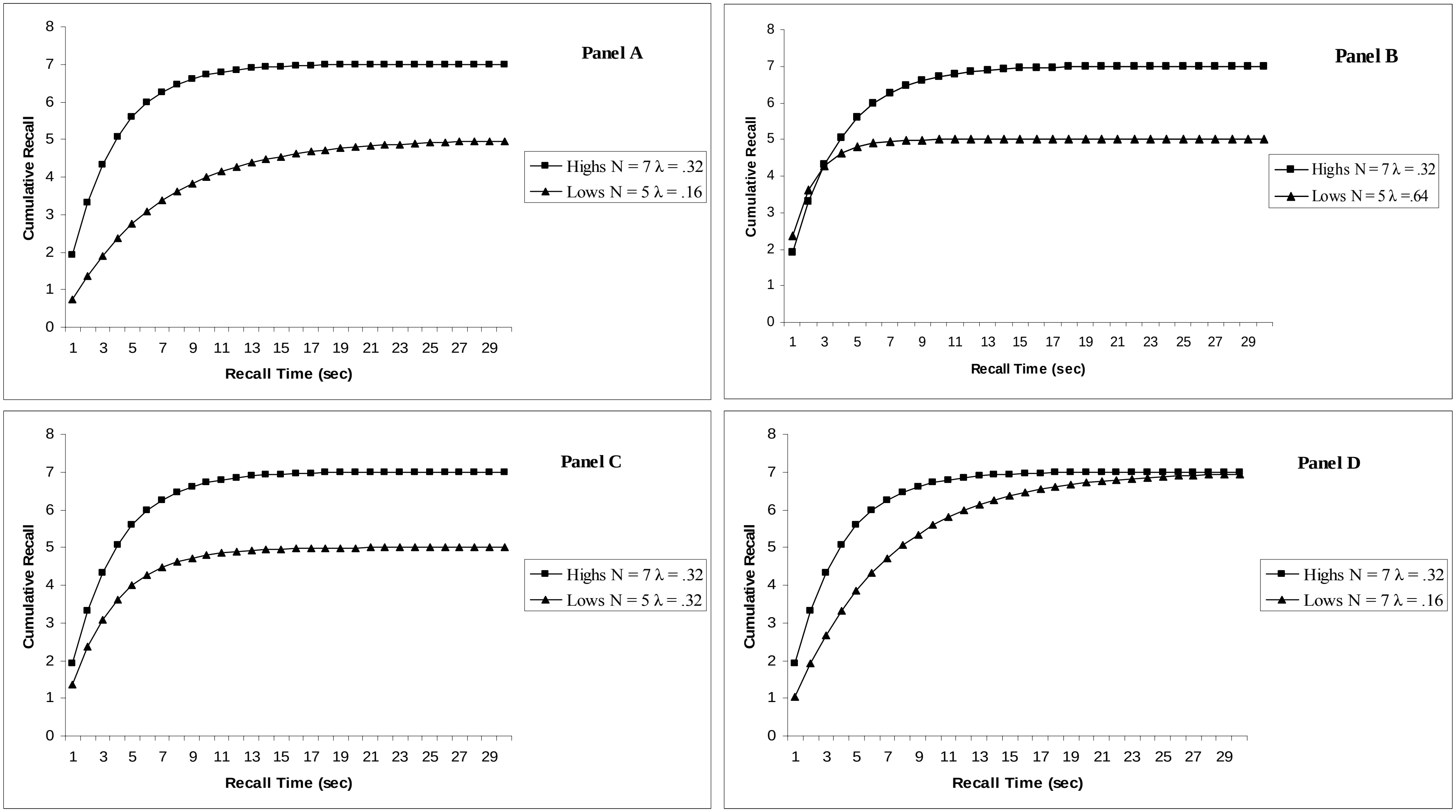

Figure 1. Hypothetical cumulative latency distributions fit by Equation 1 as a function of complex span.

recall have suggested that N and λ are inversely related (e.g., Herrmann & Chaffin, 1976; Johnson, Johnson, & Mark, 1951; Kaplan, Carvellas, & Metlay, 1969) with correlations ranging from -.48 to -.75. Although, the results from Unsworth & Engle (2005d) suggest otherwise.

A third possibility (panel C in Figure 1) is that high and low span participants search through a search set of the same size, but that the low spans' search set contains fewer recoverable targets. That is, most search models assume that items that have an absolute strength greater than some value can be recovered, but that items whose absolute strength falls below that value cannot be recovered. Therefore it is possible that low spans have more degraded (non-recoverable) targets than high spans. In some sense, this possibility examines encoding differences between high and low spans in which low spans do not sufficiently encode some items and thus these items are non-recoverable. This would result in fewer items being recalled and low values of N, but the same λ, mean recall latency, and mean IRT values because the two groups would be searching through the same size search set. A similar result has been reported by Rohrer and Wixted (1994) in terms of manipulations of presentation duration. As noted above, these authors found that increasing presentation time of items increased the number of items recalled and asymptotic levels of recall (i.e., N), but that there were no differences between the latency measures suggesting that increasing presentation rate increased the likelihood that an item would be recoverable.

The final possibility (panel D in Figure 1) is that high and low span participants search through search sets of the same size with the same number of recoverable targets but that low spans have a slower sampling time than high spans. This is a basic speed of processing account whereby low spans are slower to sample items than high spans. Thus, the reason low spans retrieve fewer items than high spans is because they are not given enough time to sample and recover all of the target items. However, given enough time, low spans should be able to recall as many items as the high spans. This would result in low spans having smaller values of λ than high spans and longer values of mean recall latency and larger mean IRTs. Crucially, however, given enough time, high and low spans should have equivalent values of N. Such a result has previously been reported by Burns and Schoff (1998) in the context of item-specific and relational processing. However, given the Unsworth and Engle (2005d) immediate free recall findings, this possibility also seems unlikely. In that study, high and low spans never reached equivalent levels of asymptotic performance, with low spans always recalling fewer items than high spans.

All four possibilities were tested by fitting the cumulative latency distributions to the cumulative exponential and examining differences in the parameter estimates. Additionally, as with the Unsworth and Engle immediate free recall study, mean recall latency and mean IRT values for both groups were examined as was mean probability correct and serial position functions.

CHAPTER 2: EXPERIMENT 1

The purpose of Experiment 1 was to examine individual differences in complex span and the dynamics of free recall in the absence of recall from primary memory. In addition, because delayed free recall has been used previously when examining the random search model (e.g., Rohrer & Wixted, 1994), Experiment 1 provides a means of replicating and extending previous findings. In terms of individual differences, each of the four possibilities presented above of recall differences between high and low complex memory span individuals was examined. The expectation was that that the differences would be most similar to the list-discrimination hypothesis (possibility 1). In order to examine these possibilities, all participants performed 21 lists with words with three different list-lengths (six, nine, or twelve items) of delayed free recall.

Method

Participant Screening for Complex Span

All participants were prescreened on three complex memory span measures. These included operation span, reading span, and symmetry span. All three tasks are computer administered and allow for group testing (e.g., Unsworth, Heitz, Schrock, & Engle, in press). The tasks have been shown to have good reliability and validity. Individuals were selected based on a z-score composite of the three tasks. Only participants falling in the upper (high spans) and lower (low spans) quartiles of the composite distribution were selected.

Operation Span (Ospan). Participants solved a series of math operations while trying to remember a set of unrelated letters (B, F, H, J, L, M, Q, R, and X). Participants were required to solve a math operation and after solving the operation they are presented with a letter for 800 ms. Immediately after the letter was presented the next operation was presented. Three trials of each list-length (3-7) were presented, with the order of list-length varying randomly. At recall, letters from the current set were recalled in the correct order by clicking on the appropriate letters (see Unsworth et al., in press for more details). Participants received three sets (of list-length two) of practice. For all of the span measures, items were scored if the item was correct and in the correct position. The score was the proportion of correct items in the correct position.

Reading Span (Rspan). Participants were required to read sentences while trying to remember the same set of unrelated letters as Ospan. For this task, participants will read a sentence and determine whether the sentence makes sense or not (e.g. "The prosecutor's dish was lost because it was not based on fact. ?"). Half of the sentences made sense while the other half did not. Nonsense sentences were made by simply changing one word (e.g. "dish" from "case") from an otherwise normal sentence. There were 10-15 words in each sentence. Participants were required to read the sentence and to indicate whether it made sense or not. After participants gave their response they were presented with a letter for 800 ms. At recall, letters from the current set were recalled in the correct order by clicking on the appropriate letters. There were three trials of each set-size with list length ranging from 3–7. The same scoring procedure as Ospan was used.

Symmetry Span (Symspan). In this task participants were required to recall sequences of red squares within a matrix while performing a symmetry-judgment task. In the symmetry-judgment task participants were shown an 8 x 8 matrix with some squares filled in black. Participants decided whether the design was symmetrical about its vertical axis. The pattern was symmetrical half of the time. Immediately after determining whether the pattern was symmetrical, participants were presented with a 4 x 4 matrix with one of the cells filled in red for 650 ms. At recall, participants recalled the sequence of red-square locations in the preceding displays, in the order they appeared by clicking on the cells of an empty matrix. There were three trials of each set-size with list length ranging from 2-5. The same scoring procedure as Ospan was used.

Composite Score

For the composite score, scores for each of the three complex span tasks were z-transformed for each participant. These z-scores were then averaged together and quartiles were computed from the averaged distribution. This distribution consisted of scores for over 600 individual participants who completed each of the three span tasks. High and low span participants in the current study were selected from this overall distribution. Additionally, participants were only selected if they maintained 80% accuracy on the processing component across the three span tasks.

Participants and Design

Participants were 25 high spans and 20 low spans, as determined by the composite measure. Participants were recruited from a subject-pool at Georgia Institute of Technology and from the Atlanta community through newspaper advertisements. Participants were between the ages of 18 and 35 and received either course credit or monetary compensation for their participation. Each participant was tested individually in a laboratory session lasting approximately one hour. Participants performed two practice lists with letters and 21 lists with words with three different list-lengths (six, nine, or twelve items).

Procedure

Participants were tested individually in the presence of an experimenter. Items were presented alone for 1 s each (similar to the immediate free recall experiment). Participants were required to read each word aloud as it appeared. After list presentation, participants engaged in a 20 s distractor task before recall: Participants saw 10 three-digit numbers appear for 2 s each, and were required to say the digits aloud in ascending order (e.g., Rohrer & Wixted, 1994). At recall participants saw three question marks appear in the middle of the screen accompanied by a brief tone indicating that the recall period had begun. Participants had 45 s to recall as many of the words as possible in any order they wished from the current trial. For each spoken response (both correct and incorrect responses) an experimenter pressed a key indicating when in the recall period the response was given.

Results

Participants

Data for two high span participants was excluded from data analyses due to data collection problems. The mean z-scores scores for the final 23 high and 20 low span participants were .98 (*SD* = .15, range .71 to 1.28) and -1.07 (*SD* = .50, range -2.36 to -.53), respectively. The mean ages for the high and low span participants were 20.43 (*SD* = 3.54) and 23.25 (*SD* = 5.95), respectively (see appendix for additional participant characteristics).

Accuracy

Probability Correct. As expected, the results suggested that classic list-length effects were apparent in which probability correct decreased as list-length increased. Additionally, as expected, high spans consistently recalled more items than low spans. These observations were confirmed by a 2 (span) by 3 (list-length) mixed analysis of variance (ANOVA), with list-length as the within-subjects variable. The ANOVA demonstrated a strong main effect of list-length, $F(2, 82) = 56.94$, $MSE = .004$, $p < .01$, partial $\eta^2 = .58$, with probability correct decreasing as list-length increased (list-length 6 $M = .58$, $SE = .02$, list-length 9 $M = .50$, $SE = .02$, list-length 12 $M = .42$, $SE = .02$). There was also a strong main effect of span, $F(1, 41) = 19.37$, $MSE = .04$, $p < .01$, partial $\eta^2 = .32$, in which high spans recalled a higher proportion of items than low spans ($M = .58$, $SE = .02$ vs. $M = .42$, $SE = .03$). Furthermore, these two factors interacted suggesting that the high span advantage was greatest at list-length six (high span: list-length 6 $M = .68$, $SE = .03$, list-length 9 $M = .57$, $SE = .03$, list-length 12 $M = .48$, $SE = .02$, low span: list-length 6 $M = .47$, $SE = .03$, list-length 9 $M = .43$, $SE = .03$, list-length 12 $M = .37$, $SE = .03$), $F(2, 82) = 6.78$, $MSE = .004$, $p < .01$, partial $\eta^2 = .14$. Thus, high spans tended to recall a higher proportion of items than low spans and this was especially true for the smaller list-lengths.

Figure 2 shows probability correct as a function of serial position for both high and low span individuals for each of the three list-lengths. Both high and low spans generated serial position functions with intact primacy and diminished recency effects consistent with prior work using delayed free recall (Glanzer & Cunitz, 1966). Additionally, span differences occurred at nearly all positions with high spans

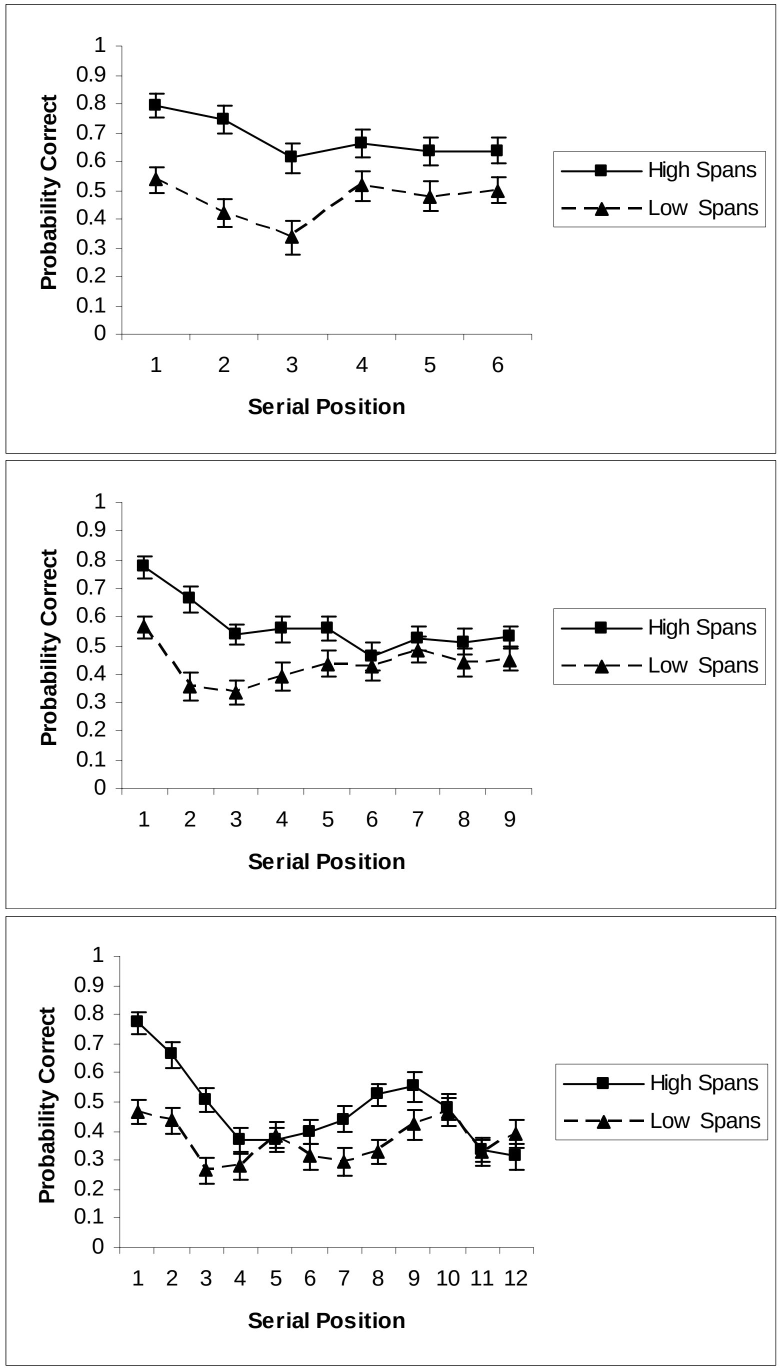

Figure 2. Probability correct as a function of serial position and complex span for each list-length in Experiment 1. Top panel = list-length 6; middle panel = list-length 9; bottom panel = list-length 12. Error bars represent one standard error of the mean.

consistently recalling more items than low spans. These observations were supported by a series of ANOVAs examining span by serial position for each of the three list-lengths. For list-length six the ANOVA demonstrated main effects of both serial position, $F(5, 205) = 5.03$, $MSE = .03$, $p < .01$, partial $\eta^2 = .11$, and span, $F(1, 41) = 23.99$, $MSE = .13$, $p < .01$, partial $\eta^2 = .37$. These two factors also interacted suggesting that the span differences were largest for the primacy portion of the curve, $F(5, 205) = 2.34$, $MSE = .03$, $p < .05$, partial $\eta^2 = .05$. Similar results were obtained for both list-length nine and list-length twelve. For list-length nine the ANOVA demonstrated main effects of both serial position, $F(8, 328) = 7.46$, $MSE = .03$, $p < .01$, partial $\eta^2 = .15$, and span, $F(1, 41) = 13.45$, $MSE = .13$, $p < .01$, partial $\eta^2 = .25$. This two way interaction was also significant, $F(8, 328) = 3.38$, $MSE = .03$, $p < .01$, partial $\eta^2 = .08$. For list-length twelve the ANOVA demonstrated main effects of both serial position, $F(11, 451) = 11.77$, $MSE = .03$, $p < .01$, partial $\eta^2 = .22$, and span, $F(1, 41) = 11.15$, $MSE = .15$, $p < .01$, partial $\eta^2 = .21$. The two way interaction was also significant, $F(11, 451) = 4.75$, $MSE = .03$, $p < .01$, partial $\eta^2 = .10$. Together, these results suggest that high spans recall a higher proportion of items at nearly all serial positions for each of the three list-lengths than low spans, with the largest span differences occurring at the primacy portion of the curve.

Recall Errors. In addition to examining probability correct, the different errors that individuals make in free recall were examined. Errors were classified as previous list intrusions (items from previous lists), extra-list intrusions (items not presented in any other list), or repetitions (items from the current list that had already been recalled). Shown in Table 1, are the average number of each error type per list (collapsed on list-length) as a function of span. The results suggested that high and low spans differ mainly in previous list intrusions (which came mainly from one list back), with low spans making many more previous list intrusions than high spans. There were also small span differences in extra-list intrusions, but no differences in repetitions.

Table 1. Mean number of each error type per list by complex span for Experiment 1.

	Error Type		
Complex Span	PLI	ELI	Repeat
High Span	.14 (.06)	.12 (.06)	.04 (.02)
Low Span	.41 (.06)	.28 (.07)	.07 (.02)

Note. PLI = previous list intrusion; ELI = extra-list intrusion; Repeat = repetition error. Numbers in parentheses are standard errors.

Separate ANOVAs examining span and list-length for each error type supported these observations. For previous list intrusions, the only significant effect was a main effect of span, $F(1, 41) = 10.59$, $MSE = .22$, $p < .01$, partial $\eta^2 = .21$. On average, these intrusions came from approximately two lists back, with the majority coming from one list back. This did not differ as a function of either list-length or span (both *p*'s > .20). For extra-list intrusions the only effect to approach conventional levels of significance was a main effect of span, $F(1, 41) = 2.87$, $MSE = .28$, $p = .10$, partial $\eta^2 = .07$. For repetitions, the only effect to approach conventional levels of significance was a main effect of list-length with the number of repetitions increasing as list-length increased, $F(2, 82) = 2.72$, $MSE = .007$, $p = .07$, partial $\eta^2 = .06$. These results suggest that high and low spans differ primarily in the number of previous list intrusions, with low spans emitting more previous list intrusions than high spans, the majority of which come from the immediately preceding list.

Latency

Cumulative Latency Distributions. Shown in Figure 3a, are the cumulative latency distributions fit to the cumulative exponential for high and low span participants (collapsed on list-length). Figure 3b, shows the cumulative latency distributions for each list-length fit to the cumulative exponential. For each, responses were first placed into forty-five one second bins, and then the cumulative number of items recalled for each bin was computed. As can be seen, the fits for each function were quite good, accounting for 98% of the variance. Furthermore, Kolmogorov-Smirnov tests examining differences between the raw and fitted values for each function resulted in non-significant *p*-values (all *p*'s > .12). As with the probability correct analyses, high spans recalled more items than low spans (i.e., higher asymptotic levels, N) and list-length effects were apparent. Additionally, as shown in Figure 3a, low spans tended to reach asymptotic levels at a slower rate (λ) than high spans.

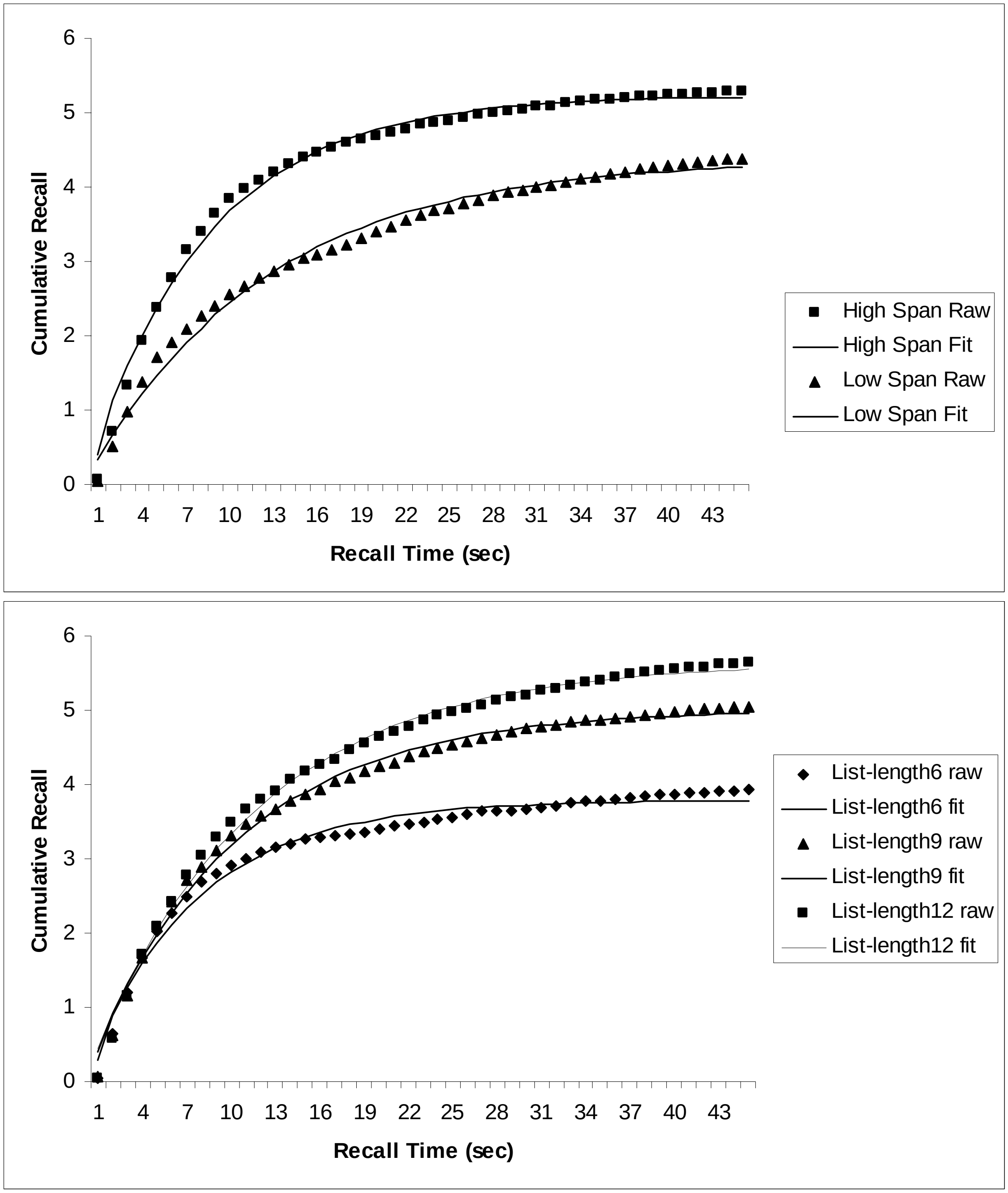

Figure 3. Cumulative recall functions for complex span and list-length in Experiment 1. The symbols represent the raw data and the solid lines represent the best fitting exponential (Equation 1).

Rate of approach to asymptote (λ) also changed as a function of list-length with rate decreasing as list-length increased consistent with Rohrer and Wixted (1994). Table 2 shows the parameter values from fitting the cumulative exponential to the cumulative latency distributions for each individual and each group for both N and λ as a function of list-length and complex span.

Table 2. Parameter estimates obtained from fitting the cumulative latency distributions to a cumulative exponential as a function of complex span and list-length for Experiment 1.

	Individual Estimates		Group Estimates	
Complex Span (List-Length)	λ	N	λ	N
Low (6)	.13	3.68	.10	3.39
High (6)	.17	4.30	.16	4.19
Low (9)	.09	4.96	.08	4.57
High (9)	.12	5.51	.12	5.48
Low (12)	.09	5.32	.07	5.24
High (12)	.11	6.22	.10	6.10

Note. λ = rate of approach to asymptotic performance; N = asymptotic performance.

In order to examine these observations, the cumulative exponential function was fit to each participant's cumulative latency distributions for each list-length. The resulting parameter estimates were then submitted to separate ANOVAs examining span and list-length. Examining first asymptotic levels of performance (N), the ANOVA demonstrated a main effect of list-length, $F(2, 80) = 41.41$, $MSE = .84$, $p < .01$, partial $\eta^2 = .51$, with N increasing as list-length increased.[1] The main effect of span approached conventional levels of significance, $F(1, 40) = 3.63$, $MSE = 4.10$, $p = .06$, partial $\eta^2 = .08$.

The two way interaction was not significant ($F < 1$). The results are generally consistent with the probability correct analyses, demonstrating list-length effects and span differences. The reason the span effect did not reach conventional levels of significance is most likely due to the large amount of variability present within the parameter estimates.

Examining next rate of approach to asymptote (λ), the ANOVA demonstrated main effects of both list-length, $F(2, 82) = 27.28$, $MSE = .001$, $p < .01$, partial $\eta^2 = .40$, and span, $F(1, 41) = 5.17$, $MSE = 007$, $p < .05$, partial $\eta^2 = .11$. The list-length effect suggested that as list-length increased, rate of approach (λ) decreased (list-length 6 $M = .15$, $SE = .01$, list-length 9 M [1]= .11, $SE = .01$, list-length 12 $M = .10$, $SE = .01$). The span effect suggested that high spans approached asymptotic levels at a faster rate than low spans ($M = .14$, $SE = .01$ vs. $M = .10$, $SE = .01$). The two way interaction was not significant ($p > .32$).

Recall Latency. In addition to examining the cumulative latency distributions and the parameter estimates from fitting the cumulative exponential, recall latency was examined. Here recall latency refers to the time point in the recall period when a given response was emitted. Thus, if responses were emitted 5, 10, and 15 s into the recall period, mean recall latency would be 10 s. Recall latency was examined partially because the parameter estimates that are derived from fitting the cumulative exponential to the cumulative latency distributions tend to be somewhat variable and can lead to low statistical power. Additionally, the cumulative latency distributions provide a fairly gross

---

[1] One low span was dropped from these analyses due to extremely large (i.e., three standard deviations above the mean) estimates of N. Including this participant in the analyses lead to qualitatively identical results.

measure of recall latency during the recall period. Therefore, mean recall latency was examined by both list-length and span. Table 3 shows mean recall latency as a function of list-length and span. The results suggested that low spans had longer recall latencies than high spans and recall latency increased with increases in list-length. These results complimented the cumulative latency distribution analyses.

Table 3. Mean latency (in seconds) by complex span and list-length for Experiment 1.

	List-length		
Complex Span	6	9	12
High Span	7.26 (.82)	8.92 (.63)	9.98 (.63)
Low Span	10.73 (.88)	12.39 (.67)	12.68 (.67)

Note. Numbers in parentheses are standard errors.

These observations were supported by a 2 (span) x 3 (list-length) mixed ANOVA, with list-length as the within subjects variable. The ANOVA yielded a main effect of list-length, $F(2, 82) = 13.60$, $MSE = 4549335$, $p < .01$, partial $\eta^2 = .25$, with recall latency increasing as list-length increased (list-length 6 $M = 9.00$ s, $SE = .60$, list-length 9 $M = 10.65$ s, $SE = .46$, list-length 12 $M = 11.33$ s, $SE = 46$). The ANOVA also yielded a main effect of span, $F(1, 41) = 13.53$, $MSE = 24546379$, $p < .01$, partial $\eta^2 = .25$, with high spans having shorter mean recall latencies than low spans ($M = 8.71$ s, $SE = .60$ vs. $M = 11.94$ s, $SE = .64$). The two way interaction was not significant ($F < 1$). Similar to the cumulative distribution analyses, the results suggested that as list-length increased recall latency increased (Rohrer & Wixted, 1994), and high spans had shorter latencies than low spans.

Next the growth of recall latencies by span was examined for the first four responses collapsed on list-length. This was done to examine differences between high and low spans in recall latency after equating for the number of responses. That is, it is possible that the reason low spans have longer recall latencies than high spans is because they recall items at the same rate as high spans initially, but that low spans are more likely than high spans to recall an item near the end of the recall period. These very late responses would then increase mean latency even though initial responses occurred at the same time. Conversely, it is possible that low spans recall items at the same rate as high spans, but they begin recalling items much later in the recall period than high spans. This would also result in longer mean latencies for lows compared to highs. Thus, in order to examine these possibilities more thoroughly, the recall latency growth functions for the first four responses were examined. Figure 4 shows the resulting recall latency growth functions for high and low spans.

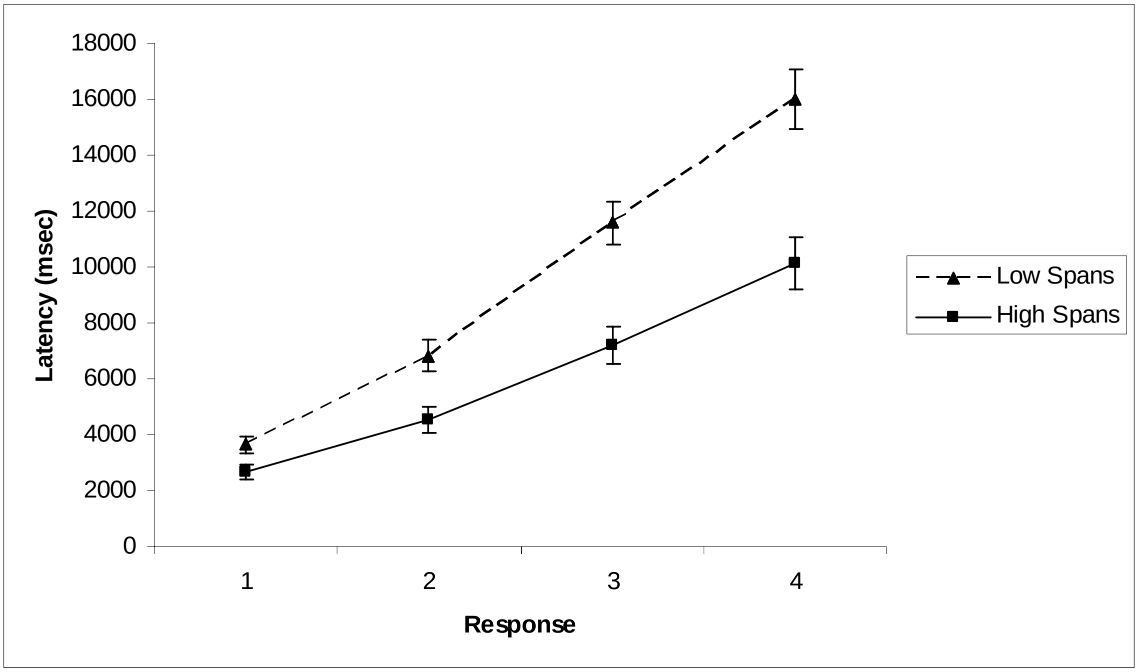

Figure 4. Recall latencies for the first four responses as a function of complex span in Experiment 1. Error bars represent one standard error of the mean.

The results suggested that high and low spans begin emitting responses at roughly the same time, but that low spans emit responses slower than high spans thereafter, resulting in overall differences in mean latency. Furthermore, these results occurred when only looking at the first four response. These notions were supported by a 2 (span) x 4 (response) mixed ANOVA, with response as the within subjects variable.[2] The ANOVA yielded main effects of response (of course), $F(3, 234) = 171.39$, $MSE = 7321374$, $p < .01$, partial $\eta^2 = .82$, and span, $F(1, 39) = 18.47$, $MSE = 74690969$, $p < .01$, partial $\eta^2 = .32$. Similar to the overall recall latency analysis, the main effect of span suggested that high spans had shorter latencies than low spans on average ($M = 6.14$ s, $SE = .52$ vs. $M = 9.52$ s, $SE = .59$). Additionally, these two factors interacted, $F(3, 234) = 10.89$, $MSE =$

[2] Two low spans were excluded from these analyses due to a low number of responses for list-length twelve.

7321374, $p < .01$, partial $\eta^2 = .22$, suggesting that high and low spans first response was emitted at about the same time, but that the span differences increased thereafter.

. *Inter-response Time.* The final analysis concerned inter-response times (IRTs). Here IRTs refer to the time between the beginning of recall of one item to recall of the next item. IRTs provide a fairly clear picture of the retrieval process. Thus, in order to examine span differences in retrieval processes, IRTs were examined. Consistent with the previous latency analyses, the results suggested that IRTs increased throughout the recall period and high spans had faster IRTs than low spans. These observations were supported initially by a 2 (span) x 3 (list-length) ANOVA, with list-length as the within subjects variable. The only statistically significant effect was a main effect of span, $F(1, 41) = 16.50$, $MSE = 7857606$, $p < .01$, partial $\eta^2 = .29$, in which high spans demonstrated faster IRTs than low spans ($M = 3.65$ s, $SE = .34$ vs. $M = 5.66$ s, $SE = .36$). This effect compliments the recall latency analyses suggesting that low spans emit responses slower than high spans leading to slower IRTs.

Similar to the recall latency growth function analyses, IRT growth functions were examined for the first three IRTs collapsed on list-length. As with those analyses, two low span participants were excluded from the analysis. The results suggested that IRTs increased throughout the recall period and high spans had faster IRTs than low spans even for only three IRT values. A 2 (span) x 3 (IRT interval) mixed ANOVA, with IRT interval as the within subjects variable supported these notions. The ANOVA yielded a main effect of IRT interval, $F(2, 156) = 10.12$, $MSE = 7715832$, $p < .01$, partial $\eta^2 = .21$, in which IRTs increased throughout the recall period (IRT interval 1-2 $M = 3.09$ s, $SE = .21$, IRT interval 2-3 $M = 3.86$ s, $SE = .28$, IRT interval 3-4 $M = 4.92$ s, $SE = .43$). There was also a main effect of span, $F(1, 39) = 21.89$, $MSE = 16759904$, $p < .01$, partial $\eta^2 = .36$, with high spans demonstrating faster mean IRTs than low spans ($M = 2.95$ s, $SE = .28$ vs. $M = 4.96$ s, $SE = .32$) for the first three IRTs. Thus, as with the recall latency analysis, the time it took to emit responses increased throughout the recall period and low spans were slower to emit responses than high spans.

Discussion

The results from Experiment 1 were largely consistent with the predictions for the list-discrimination hypothesis and inconsistent with the other three possibilities. According to the list-discrimination hypothesis, the reason low spans recall fewer items than high spans is because they do not effectively delimit their search to only current target items, but also include items from previous trials. That is, they are poorer at discriminating items belonging to list N from items belonging to list N – 1. This failure in list-discrimination leads to many previous list intrusions being included in low spans' search sets which leads to greater cue-overload and a lower overall probability of correct recall. Accordingly, the list-discrimination view suggests that low spans should recall fewer items than high spans and recall more previous list intrusions than high spans. Both predictions were supported by the data. Low spans consistently recalled fewer items than high spans and recalled more previous list intrusions than high spans. Additionally, these previous list intrusions tended to come from the immediately preceding list supporting the list discrimination view.

The list discrimination view also predicted that because low spans have more intruding items in their search sets, they should be slower than high spans to sample and emit items from secondary memory. That is, according to the random search model the more items within the search set, the longer it should take on average to find a desired item. Consequently, because low spans have larger search sets than high spans, they should recall items at a slower rate than high spans leading to overall differences in recall latency and IRTs. These predictions were also supported by the data. The latency analyses suggested that low spans approached asymptotic levels of performance at a slower rate than high spans, recalled responses at a slower rate than high spans, and took longer in between the recall of items than high spans. According to the random search model, these results suggest that low spans are searching through a larger set of items than high spans (e.g., Rohrer & Wixted, 1994). Indeed, examining Tables 2 and 3 suggests that low spans are searching through roughly the same number of items at a list-length of six as high spans are at a list-length of twelve. Thus, even at lower list-lengths, low spans search sets are much larger than those of high spans. Together, these results suggest that low spans search through a larger set of items than high spans which leads to slower and less accurate recall than high spans.

CHAPTER 3: EXPERIMENT 2

In Experiment 2 span differences in episodic retrieval and the dynamics of free recall were examined in the continuous distractor free recall paradigm (Bjork & Whitten, 1974). The reasons for examining performance on the continuous distractor paradigm were threefold. 1) Continuous distractor free recall is very similar to the design of the complex working memory span tasks, differing only in type of recall (free vs. serial) and type of retention interval (filled vs. unfilled). That is, in complex span tasks serial recall is required and there is typically no distracting task after the lasted presented item. Accordingly, examining the dynamics of free recall in the continuous distractor task should provide a fairly accurate portrayal of retrieval in the complex working memory span tasks. 2) No previous study has fully examined the time-course of retrieval in the continuous distractor task and the possibility of important individual differences therein. 3) The results from continuous distractor free recall should replicate the basic pattern of results obtained with delayed free recall.

As with the original continuous distractor paradigm, to-be-remembered items were interspersed with a distracting activity. Interspersed between each to-be-remembered item participants were required to arrange a series of three digit numbers in ascending order (the same distracting task as Experiment 1). After the presentation of the last item, participants engaged in an additional 16 s of distractor activity during the retention interval. The hypotheses and analyses for Experiment 2 were exactly the same as those for Experiment 1.

Participants and Design

Participants were 40 new high (n = 20) and low (n = 20) spans, as determined by the composite measure. Participant recruitment and prescreening on complex memory span was exactly the same as Experiment 1. Each participant was tested individually in a laboratory session lasting approximately one hour. Participants performed two practice lists with letters and 21 lists with words with three different list-lengths (six, nine, or twelve items).

Procedure

Participants were tested one a time in the presence of an experimenter. Items were presented alone for 1 s each. Participants were required to read each word aloud as it appeared. Before and after each item presentation participants were required to arrange four separate three digit numbers (presented for 2 s each) in ascending order aloud. After list presentation, participants engaged in an additional 16 s distractor activity (e.g., eight three digit numbers instead of four) before recall. At recall participants saw three question marks appear in the middle of the screen accompanied by a tone that indicated that the recall period had begun. Participants had 45 s to recall as many of the words as possible in any order they wish. For each spoken response (both correct and incorrect responses) an experimenter pressed a key indicating when in the recall period the response was given.

Results

Participants

The mean z-scores scores for the 20 high and 20 low span participants were .96 (*SD* = .21, range .71 to 1.39) and -1.11 (*SD* = .56, range -2.96 to -.52), respectively. The mean ages for the high and low span participants were 22.10 (*SD* = 4.94) and 25.45 (*SD* = 5.67), respectively (see appendix for additional participant characteristics).

Accuracy

Probability Correct. As with Experiment 1, classic list-length effects were apparent in which probability correct decreased as list-length increased and high spans consistently recalled more items than low spans. A 2 (span) by 3 (list-length) mixed ANOVA, with list-length as the within-subjects variable supported these observations. The ANOVA yielded a main effect of list-length, $F(2, 76) = 89.00$, $MSE = .004$, $p < .01$, partial $\eta^2 = .70$, with probability correct decreasing as list-length increased (list-length 6 $M = .74$, $SE = .02$, list-length 9 $M = .61$, $SE = .02$, list-length 12 $M = .57$, $SE = .02$). There was also a main effect of span, $F(1, 38) = 28.66$, $MSE = .06$, $p < .01$, partial $\eta^2 = .43$, in which high spans recalled a higher proportion of items than low spans ($M = .76$, $SE = .03$ vs. $M = .52$, $SE = .03$). The two way interaction was not statistically significant ($F < 1$). Thus, probability correct decreased as list-length increased and high spans tended to recall a higher proportion of items than low spans.

Shown in Figure 5 is probability correct as a function of serial position for both high and low span individuals for each of the three list-lengths. As expected, both high and low spans generated serial position functions with intact primacy and recency effects consistent with prior work using continuous distractor free recall (Bjork & Whitten, 1974). Additionally, span differences occurred at all positions with high spans consistently recalling more items than low spans.

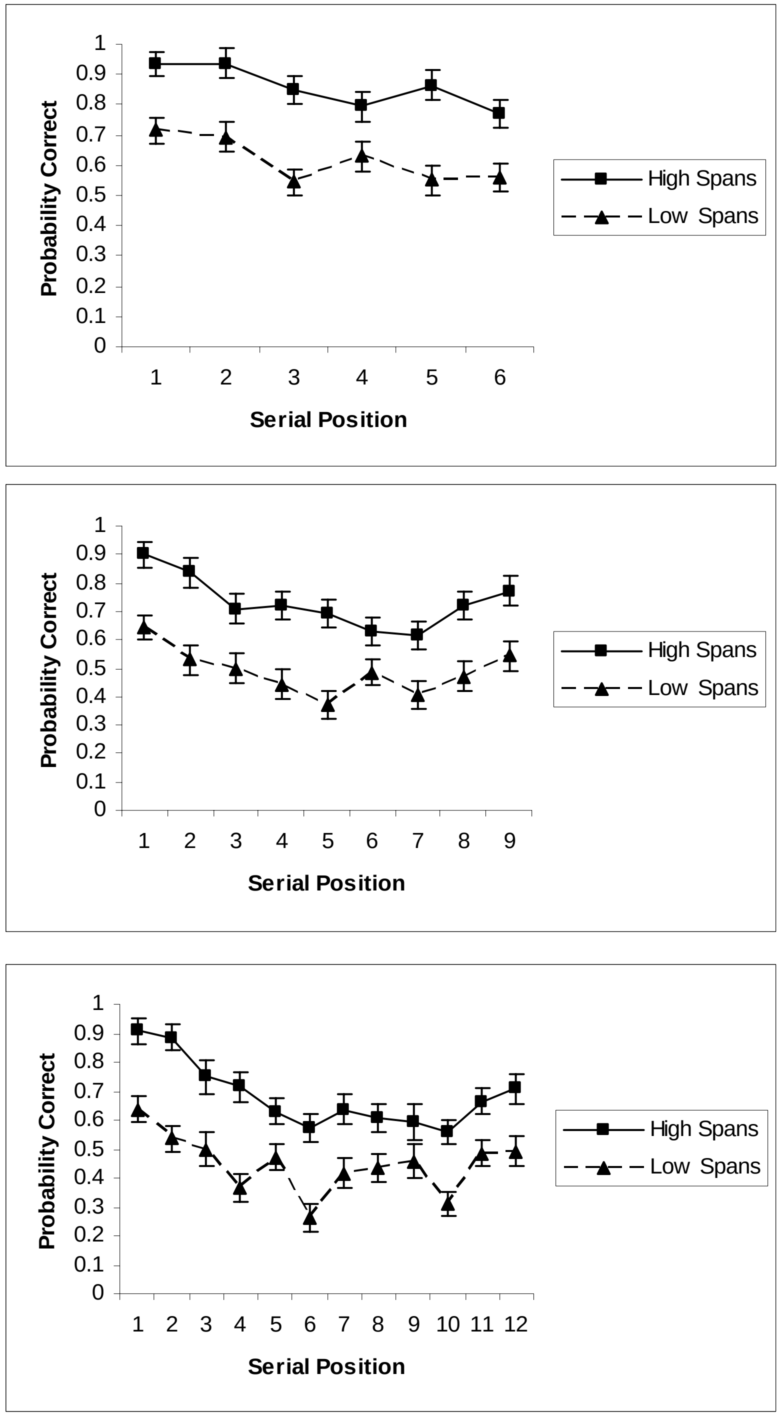

Figure 5. Probability correct as a function of serial position and complex span for each list-length in Experiment 2. Top panel = list-length 6; middle panel = list-length 9; bottom panel = list-length 12. Error bars represent one standard error of the mean.

These observations were supported by a series of ANOVAs examining span by serial position for each of the three list-lengths. The ANOVA for list-length six demonstrated main effects of both serial position, $F(5, 190) = 7.03$, $MSE = .03$, $p < .01$, partial $\eta^2 = .16$, and span, $F(1, 38) = 25.11$, $MSE = .14$, $p < .01$, partial $\eta^2 = .40$. These two factors did not interact ($p > .25$). Similar results were obtained for both list-length nine and list-length twelve. For list-length nine the ANOVA demonstrated main effects of both serial position, $F(8, 304) = 8.36$, $MSE = .03$, $p < .01$, partial $\eta^2 = .18$, and span, $F(1, 38) = 26.68$, $MSE = .20$, $p < .01$, partial $\eta^2 = .41$. This two way interaction was not significant ($F < 1$). For list-length twelve the ANOVA demonstrated main effects of both serial position, $F(11, 418) = 15.66$, $MSE = .03$, $p < .01$, partial $\eta^2 = .29$, and span, $F(1, 38) = 23.45$, $MSE = .29$, $p < .01$, partial $\eta^2 = .38$. Here, the two way interaction was also significant, $F(11, 418) = 1.96$, $MSE = .03$, $p < .05$, partial $\eta^2 = .05$. The interaction suggested that the span differences were largest for the primacy portion of the serial position function. Together, the results suggested that high spans recall a higher proportion of items at all serial positions for each of the three list-lengths than low spans, with slightly larger span differences occurring at the primacy portion of the curve.

Recall Errors. As with Experiment 1, recall errors were also examined. These errors were classified as previous list intrusions (items from previous lists), extra-list intrusions (items not presented in any other list), or repetitions (items from the current list that had already been recalled). Table 4 shows the average number of each error type per list (collapsed on list-length) as a function of span. Similar to Experiment 1 the results suggested that high and low spans differ in previous list intrusions (which came mainly from one list back), with low spans making many more previous list intrusions than high spans. Span differences did not occur for either extra-list intrusions or repetitions.

Table 4. Mean number of each error type per list by complex span for Experiment 2.

	Error Type		
Complex Span	PLI	ELI	Repeat
High Span	.05 (.10)	.06 (.06)	.06 (.03)
Low Span	.43 (.10)	.17 (.06)	.11 (.03)

Note. PLI = previous list intrusion; ELI = extra-list intrusion; Repeat = repetition error. Numbers in parentheses are standard errors.

Separate ANOVAs examining span and list-length for each error type supported these observations. For previous list intrusions, the only significant effect was a main effect of span, $F(1, 38) = 7.64$, $MSE = .58$, $p < .01$, partial $\eta^2 = .17$. Consistent with delayed free recall these intrusions, on average, came from approximately two lists back, with the majority coming from one list back. This did not differ as a function of either list-length or span (both *p*'s > .20). For extra-list intrusions none of the effects reached conventional levels of significance (all *p's* > .11). For repetitions, the only effect to reach conventional levels of significance was a main effect of list-length with the number of repetitions increasing as list-length increased, $F(2, 76) = 6.11$, $MSE = .01$, $p < .01$, partial $\eta^2 = .14$. Together the results suggested that high and low spans differ only in the number of previous list intrusions they make, with low spans emitting more previous list intrusions than high spans, the majority of which come from the immediately preceding list.

Latency

Cumulative Latency Distributions. Figure 6a, shows the cumulative latency distributions fit to the cumulative exponential for high and low span participants (collapsed on list-length). Shown in Figure 6b, are the cumulative latency distributions for each list-length fit to the cumulative exponential. As with Experiment 1 responses were first placed into forty-five one second bins, and then the cumulative number of items recalled for each bin was computed. As can be seen, the fits for each function were quite good, accounting for 98% of the variance. Furthermore, Kolmogorov-Smirnov tests examining differences between the raw and fitted values for each function resulted in non-significant *p*-values for all functions (all *p*'s > .43). Similar to the probability correct analyses, high spans recalled more items than low spans (i.e., higher asymptotic levels, N) and list-length effects were apparent. Furthermore, as shown in Figure 6a, and consistent with Experiment 1, low spans tended to reach asymptotic levels at a slower rate (λ) than high spans.

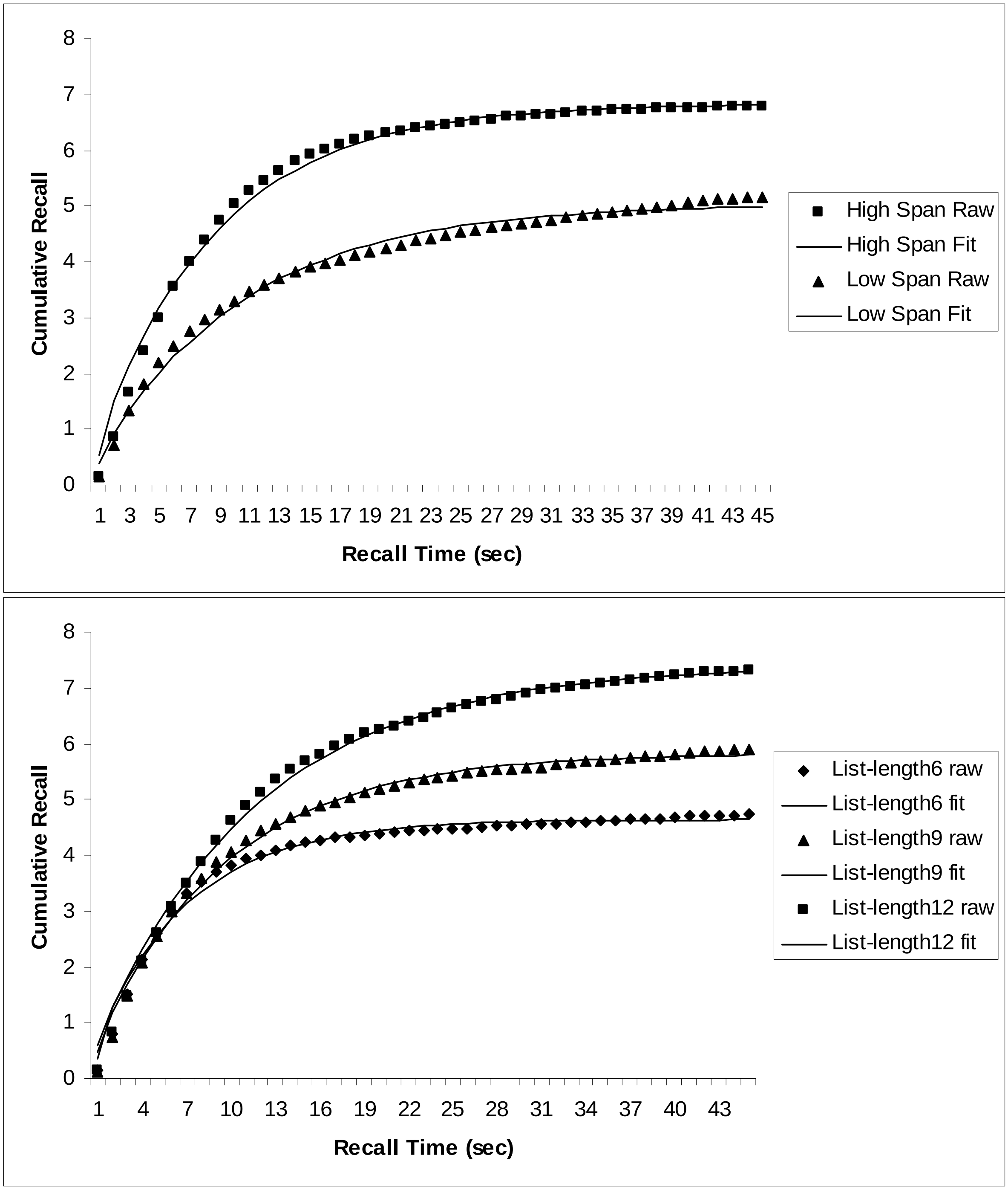

Figure 6. Cumulative recall functions for complex span and list-length in Experiment 2. The symbols represent the raw data and the solid lines represent the best fitting exponential (Equation 1).

Rate of approach to asymptote (λ) also changed as function of list-length with rate decreasing as list-length increased consistent with Experiment 1 and the work of Rohrer and Wixted (1994). Table 5 shows the parameter values from fitting the cumulative exponential to the cumulative latency distributions for each individual and each group for both N and λ as a function of list-length and complex span.

Table 5. Parameter estimates obtained from fitting the cumulative latency distributions to a cumulative exponential as a function of complex span and list-length for Experiment 2.

	Individual Estimates		Group Estimates	
Complex Span (List-Length)	λ	N	λ	N
Low (6)	.15	4.23	.13	4.08
High (6)	.19	5.25	.18	5.24
Low (9)	.12	4.81	.10	4.91
High (9)	.14	6.85	.12	6.78
Low (12)	.09	6.20	.08	6.23
High (12)	.11	8.78	.10	8.64

Note. λ = rate of approach to asymptotic performance; N = asymptotic performance.

Similar to Experiment 1, the cumulative exponential function was fit to each participant's cumulative latency distributions for each list-length. The resulting parameter estimates were then submitted to separate ANOVAs examining span and list-length. The ANOVA examining asymptotic levels of performance (N) yielded a main effect of list-length, $F(2, 76) = 59.67$, $MSE = 1.29$, $p < .01$, partial $\eta^2 = .61$, with N increasing as list-length increased. The main effect of span was also significant, $F(1, 38) = 11.47$, $MSE = 4.10$, $p < .01$, partial $\eta^2 = .23$. Furthermore, these two factors interacted,

$F(2, 76) = 4.87$, $MSE = 1.29$, $p < .01$, partial $\eta^2 = .11$, with the span differences increasing as list-length increased. These results are generally consistent with the probability correct analyses, demonstrating list-length effects and span differences.

The ANOVA examining rate of approach to asymptote (λ) demonstrated a main effect of list-length, $F(2, 74) = 56.61$, $MSE = .001$, $p < .01$, partial $\eta^2 = .61$.[3] The list-length effect suggested that as list-length increased, rate of approach (λ) decreased (list-length 6 $M = .15$, $SE = .01$, list-length 9 $M = .11$, $SE = .01$, list-length 12 $M = .10$, $SE = .01$). The main effect of span approached conventional levels of significance, $F(1, 37) = 3.51$, $MSE = 007$, $p = .07$, partial $\eta^2 = .09$. The span effect suggested that high spans approached asymptotic levels at a faster rate than low spans ($M = .15$, $SE = .01$ vs. $M = .12$, $SE = .01$). The two way interaction was not significant ($p > .22$).

Recall Latency. As with Experiment 1, recall latency was also examined to get a better picture on the time taken to emit responses. As mentioned previously, these analyses have greater power than the analyses based on the parameter estimates. Again, recall latency refers to the time point in the recall period when a given response was emitted. Shown in Table 6 is mean recall latency as a function of list-length and span. Consistent with Experiment 1 and the cumulative latency distribution analyses, the results suggested that low spans had longer recall latencies than high spans and recall latency increased with increases in list-length.

[3] One low span was dropped from these analyses due to extremely large estimates of λ. Including this participant in the analyses lead to qualitatively identical results.

Table 6. Mean latency (in seconds) by complex span and list-length for Experiment 2.

	List-length		
Complex Span	6	9	12
High Span	5.77 (.64)	8.02 (.70)	9.09 (.64)
Low Span	8.89 (.64)	10.98 (.70)	12.03 (.64)

Note. Numbers in parentheses are standard errors.

These observations were supported by a 2 (span) x 3 (list-length) mixed ANOVA, with list-length as the within subjects variable. The ANOVA yielded a main effect of list-length, $F(2, 76) = 47.68$, $MSE = 2281961$, $p < .01$, partial $\eta^2 = .56$, with recall latency increasing as list-length increased (list-length 6 $M = 7.33$ s, $SE = .45$, list-length 9 $M = 9.5$ s, $SE = .49$, list-length 12 $M = 10.56$ s, $SE = 45$). The ANOVA also yielded a main effect of span, $F(1, 38) = 12.61$, $MSE = 21496254$, $p < .01$, partial $\eta^2 = .25$, in which high spans had shorter mean recall latencies than low spans ($M = 7.63$ s, $SE = .60$ vs. $M = 10.63$ s, $SE = .60$). The two way interaction was not significant ($F < 1$). Similar to the cumulative distribution analyses and the results from Experiment 1, the results suggested that as list-length increased recall latency increased and high spans had shorter recall latencies than low spans.

Consistent with Experiment 1, the growth of recall latencies by span was examined for the first four responses collapsed on list-length. Shown in Figure 7 are the resulting recall latency growth functions for high and low spans.

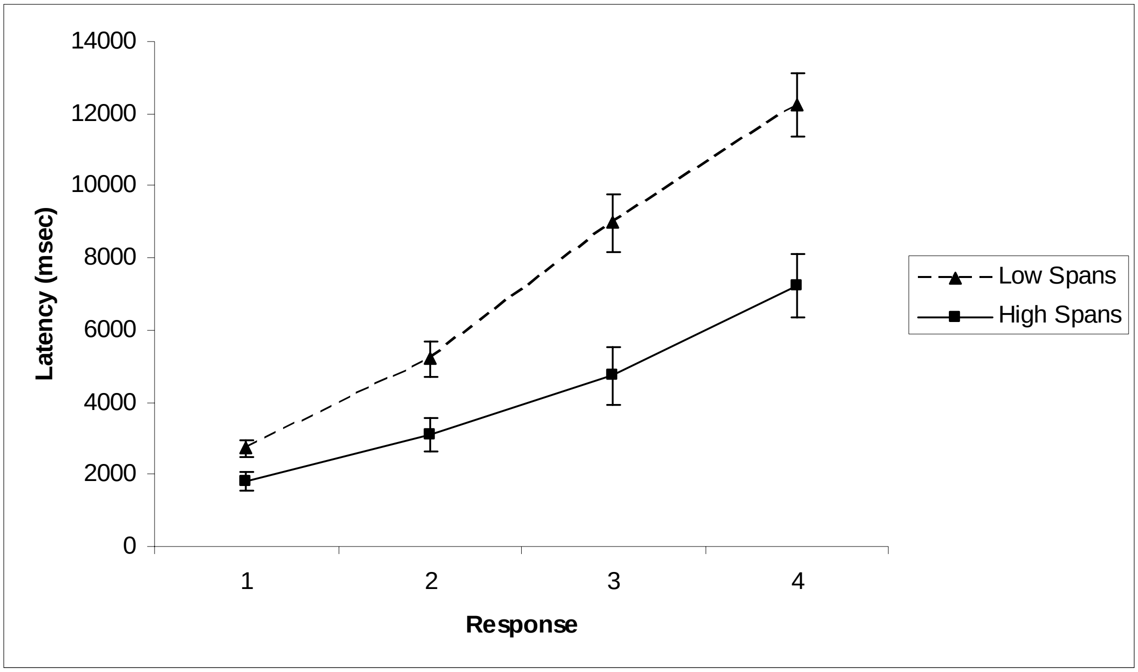

Figure 7. Recall latencies for the first four responses as a function of complex span in Experiment 2. Error bars represent one standard error of the mean.

The results suggested that high and low spans begin emitting responses at roughly the same time, but that low spans emitted responses at a slower rate than high spans thereafter, resulting in overall differences in mean latency. These observations were supported by a 2 (span) x 4 (response) mixed ANOVA, with response as the within subjects variable. The ANOVA yielded main effects of response (of course), $F(3, 228) = 169.55$, $MSE = 4209496$, $p < .01$, partial $\eta^2 = .82$, and span, $F(1, 38) = 14.55$, $MSE = 77427668$, $p < .01$, partial $\eta^2 = .28$. The main effect of span suggested that when examining only the first four responses, high spans still had shorter latencies than low spans on average ($M = 4.21$ s, $SE = .58$ vs. $M = 7.28$ s, $SE = .58$). Additionally, these two factors interacted, $F(3, 228) = 14.40$, $MSE = 4209496$, $p < .01$, partial $\eta^2 = .28$, suggesting that high and low spans first response was emitted at about the same time, but that the span differences increased thereafter.

Inter-response Time. As with Experiment 1, IRTs were examined to compliment the response latency analyses. Once again, IRTs refer to the time between the beginning of recall of one item to recall of the next item. Consistent with the previous latency analyses and Experiment 1, the results suggested that IRTs increased throughout the recall period and high spans had faster IRTs than low spans. These observations were supported by a 2 (span) x 3 (list-length) ANOVA, with list-length as the within subjects variable. The only effect to reach conventional levels of significance was a main effect of span, $F(1, 38) = 27.96$, $MSE = 5899990$, $p < .01$, partial $\eta^2 = .42$, in which high spans had faster mean IRTs than low spans ($M = 2.62$ s, $SE = .31$ vs. $M = 4.97$ s, $SE = .31$). This effect compliments the recall latency analysis suggesting that low spans emit responses slower than high spans leading to slower IRTs.

Similar to Experiment 1, IRT growth functions were examined for the first three IRTs collapsed on list-length. Complimenting the results suggested that IRTs increased throughout the recall period and, even for only the first three IRTs, high spans had faster IRTs than low. A 2 (span) x 3 (IRT interval) mixed ANOVA, with IRT interval as the within subjects variable supported these notions. The ANOVA yielded a main effect of IRT interval, $F(2, 152) = 8.13$, $MSE = 2158253$, $p < .01$, partial $\eta^2 = .18$, in which IRTs increased throughout the recall period (IRT interval 1-2 $M = 2.43$ s, $SE = .26$, IRT interval 2-3 $M = 2.66$ s, $SE = .32$, IRT interval 3-4 $M = 3.36$ s, $SE = .25$). There was also a main effect of span, $F(1, 38) = 15.78$, $MSE = 20691556$, $p < .01$, partial $\eta^2 = .29$, with high spans demonstrating faster mean IRTs than low spans ($M = 1.86$ s, $SE = .34$ vs. $M = 3.77$ s, $SE = .34$) for the first three IRTs. Thus, as with the recall latency analysis, the time it took to emit responses increased throughout the recall period and low spans were slower to emit responses than high spans.

Discussion

As with Experiment 1, the results are largely consistent with the predictions of the list discrimination view and inconsistent with the other possibilities presented in Figure 1. That is, the results suggested that low spans recalled fewer items than high spans and recalled more previous list intrusions than high spans. As with previous work these intrusions tended to come predominantly from one list back. Both results are consistent with the view that low span individuals are poorer at focusing their search on current target items than high spans and instead of discriminating between lists, low spans' search sets contain a combination of both current and previous target items. The latency analyses supported these overall notions by demonstrating that low spans reached lower asymptotic levels of recall than high spans and their rate of approach was slower than that of high spans. Accordingly, low spans recalled items at a slower rate than high spans and the time in between recalls (IRTs) was slower for low spans than for high spans. Together, these results support the contention that low spans search through a larger set of items, which includes both current and previous targets, than high spans. Indeed, consistent with Experiment 1, examining recall latency by list-length for high and low spans suggested that low spans were searching through approximately the same number of items for list-length six as high spans were for list-length twelve (see Table 6). The end result is that low spans suffer from more cue-overload than high spans resulting in slower and less accurate recall.

CHAPTER 4: GENERAL DISCUSSION

In two experiments individual differences in complex span and episodic retrieval were investigated using versions of delayed and continuous distractor free recall. Across the two experiments, it was shown that individuals who score low on complex span measures recall fewer items in free recall tasks, make more intrusions from previous lists, and recall at a slower rate than individuals who score high on complex span measures. Experiment 1 demonstrated that high and low spans differed in both overall probability correct with high spans recalling a higher proportion of items than low spans. Additionally, low spans were more likely to recall target items from previous trials than high spans and these items typically came from one list back. Furthermore, high and low spans also differed in recall latency. In all cases, low spans recalled items at a slower rate than high spans. Initially, high and low spans began recalling items at roughly the same time, but low spans took longer than high spans to recall items thereafter leading to larger IRTs. These differences in recall latency occurred even though low spans recalled consistently fewer items than high spans. That is, low spans recalled fewer items than high spans, but it took them longer than high span to recall those items. Together, the results suggested that low spans' retrieval deficits are partially due to the fact that they are searching through a larger set of items than high spans. Indeed, an examination of recall latency differences by list-length suggested that low spans search through approximately the same number of items at list-length six as high spans do at list-length twelve. This relative increase in the size of the search set leads to both less accurate and slower recall of items. Experiment 2, using continuous distractor free recall corroborated these basic findings. Low spans recalled fewer items than high spans at all list-lengths, recalled more previous list intrusions than high spans, and recalled items at a slower rate than high spans. As with the results from Experiment 1, an examination of recall latency by list-length for highs and lows suggested that low spans search through a similar number of items at list-length six as highs do at list-length twelve.

Collectively, these results provide strong support for the list-discrimination view and do not provide support for the other three possibilities. Recall that the list-discrimination view (possibility 1 in Figure 1) predicts that low spans should recall fewer items than high spans, recall more previous list intrusions than high spans, and importantly have longer recall latencies than high spans. Each of these predictions were supported by the data. The second possibility considered was that low spans' recall deficits are due to searching through a smaller set of items than high spans. This view predicts that low spans should recall fewer items than high spans and crucially, should actually have shorter recall latencies than high spans. Clearly this is not the case and possibility two is not supported by the data. The third possibility was that highs and lows search through the same number of items, but that low spans have less recoverable targets than high spans, possibly due to differences in attention at encoding. This view predicts that low spans should recall fewer items than high spans, but that both groups should have equivalent recall latencies. As with possibility two, the results do not support this view given the large differences between highs and lows in recall latency. The final possibility considered was that highs and lows search through the same number of items, have the same number of recoverable targets, but that low spans have slower sampling rates than high spans. This is a fairly simple speed of processing view that suggests high and low spans differ in the speed with which they can sample and recall items. This view predicts that high and low spans should differ in recall latency, consistent with the results, but that given enough time highs and lows should recall the same number of items. Thus, this view predicts that the reason that low spans do not recall the same number of items as high spans is because in typical recall tasks they are simply not given enough time to sample and recall all the items. Clearly the accuracy results do not support this view. Low spans consistently recalled fewer items than high spans in all conditions. Furthermore, the parameter estimates obtained from fitting the cumulative exponential to the cumulative latency distributions demonstrated that high and low spans have different levels of asymptotic performance (N). That is, N provides an index of asymptotic performance given infinite time. Thus, the differences between highs and lows in N suggest that given even a very long recall period, low spans would likely not reach the same levels of recall as high spans. Therefore, this possibility is, also, not supported by the data. Rather the data support the list-discrimination view suggesting that low spans are poorer than high spans at using temporal-contextual cues to focus search only on the current list items.

Individual Differences in WMC as Differences in Primary and Secondary Memory

Recently, Unsworth and Engle (2005d) have advanced a model of individual differences in WMC based on differences in the ability to maintain representations in primary memory via continued allocation of attention and differences in the ability to guide a cue-dependent search and retrieval process from secondary memory. Similar to the Search of Associative Memory model (SAM) advanced by Raaijmakers and Shiffrin (1980; 1981) and the context-activation model of Davelaar and colleagues (Davelaar et al., 2005), this view combines a flexible attentional component with a cue-dependent search process. In this framework, the attentional component (primary memory) serves to actively maintain a few distinct representations for on-line processing. These representations include things such as goal states for the current task, action plans, and item representations in list memory tasks. In this view, as long as attention is allocated to these representations they will be actively maintained. This continued allocation of attention serves to protect these representations from interfering information such as PI. However, if attention is removed from the representations due to internal or external distraction or due to the processing of incoming information, these representations will no longer be maintained in primary memory and if needed, will have to be retrieved from secondary memory. Accordingly, secondary memory relies on a cue-dependent search mechanism to retrieve items back into primary memory. In many cases, search of secondary memory will rely on temporal-contextual cues to reactivate items. Thus, items that have recently been displaced from primary memory will compete for retrieval with other recently displaced items. The key to successful retrieval, then, is the ability to correctly differentiate target representations from other recently presented nontarget representations. The use of precise temporal-contextual cues aids in this regard. The extent to which the search of secondary memory can be focused only on recently presented target representations, the higher will be the likelihood of retrieving the desired information.

Unsworth and Engle (2005d) argued that individuals will differ in both the ability to maintain items in primary memory and the ability to retrieve items from secondary memory. Furthermore, Unsworth and Engle argued that it is these abilities that are primarily indexed by individual differences in complex span measures. Examining high and low span differences in immediate free recall Unsworth and Engle found that high and low spans differ in both the ability to maintain items in primary memory and in the ability to retrieve items from secondary memory (see also Unsworth & Engle, in press). In terms of individual differences in secondary memory retrieval abilities, Unsworth and Engle argued that low spans used noisier temporal-contextual cues than high spans which lead to an increase in the size of low spans' search sets (cue-overload) due to an increase in the number of previous list intrusions being included in the search set. In support of this, Unsworth and Engle found that low spans recalled fewer items than high spans overall and low spans recalled many more previous list items than high spans in immediate free recall. Additionally, an examination of recall latencies suggested that low spans took longer to recall items than spans consistent with the notion that they were searching through a larger set of items than high spans.

As mentioned previously, however, these conclusions were hampered by the fact that some items were unloaded from primary memory and, thus the results do not clearly implicate retrieval differences from secondary memory. Using delayed and continuous distractor free recall alleviates this problem. Specifically, having participants engage in a distracting task just before the recall period should have displaced all items from primary memory ensuring that all subsequent retrievals would come from secondary memory. Consequently, any span differences that ensue can be interpreted as occurring due to differences in retrieval from secondary memory. With this in mind, the results suggested that, like the immediate free recall findings, low spans were both slower and less accurate at recalling items than high spans. As noted previously these results supported the list-discrimination view suggesting that low spans were using noisier temporal-contextual cues than high spans and, thus were unable to correctly distinguish current target items from previous target items leading to increased search set. Both experiments were in general agreement with the framework advocated by Unsworth and Engle (2005d) suggesting that individual differences in WMC are partially due to differences in the ability to use cues (particularly temporal-contextual cues) to guide the search process of secondary memory. In all cases, low spans suffered from greater cue-overload than high spans leading to low spans' recall deficits.

This view suggests that individual differences in many cognitive abilities, including WMC, can be thought of as occurring due to differences in the functioning of a fairly general memory system that combines a flexible attention component with a cue-dependent search mechanism. Together, these two components allow for many dynamic and complex behaviors where information must be either maintained or quickly and accurately retrieved in order for on-line processing of information to occur. These behaviors include reading a difficult passage or attempting to solve a complex matrix reasoning problem. In both cases, attention must be maintained on current information while recently presented information must also be called upon in order for integration of information.

It should be noted that this view and the present results are consistent with other views of individual differences in WMC. For instance, both the inhibition view of Hasher, Zacks, and colleagues (Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999) and the executive attention view of Conway, Engle, Kane, and colleagues (Engle & Kane, 2004; Kane, Conway, Hambrick, & Engle, in press) would be able to account for some of the present findings. For instance, both views would likely predict that low WMC individuals would be more susceptible to interference from previous trials and that this interference would result in recall problems on current trials (e.g., Kane & Engle, 2000; May, Hasher, & Kane, 1999). However, instead of arguing for differences in cue-dependent search, both views would likely suggest that low WMC individuals are more susceptible to interference from previous trials due to an inability to adequately suppress irrelevant information. Thus, according to these views, the reason low WMC individuals would search through a larger search set of items than high spans is because they are unable to suppress previous target items. Similar to Anderson and Spellman (1995), these views suggest that attention is focused internally on target representations while irrelevant representations are suppressed. Based on this notion of active suppression, both views would likely be able to handle the present results.

Additionally, it is possible that attention control may exert its influence on retrieval without recourse to a suppression mechanism. For instance, in cue-dependent models of retrieval, attention is needed for deciding how to search, what cues should be used for the search, how the cues should be combined for search, and what rules will determine when to stop the search (e.g., Raaijmakers & Shiffrin, 1980; Shiffrin, 1970). Indeed, work by Naveh-Benjamin and Guez (2000) has suggested that dividing attention at retrieval selectively disrupts the cue-elaboration aspect of the search process. That is, Naveh-Benjamin and Guez argued that attention is needed to combine cues and use to those cues to guide the search process. Thus, if attentional processes are disrupted at retrieval, cues will not be effectively combined and the search process will be more disorganized. According to this view, low spans are poorer at using attention than high spans to combine cues to delimit the search set and to use those to guide the search process. Future work is, therefore, needed to examine differences between views that rely on inhibitory processes and views that rely on cues to correctly delimit the search set.

Comparisons Between Immediate, Delayed, and Continuous Distractor Free Recall

The above framework suggesting that individual differences in WMC can be seen as arising from either primary or secondary memory provides a means to examine theoretical differences between immediate, delayed, and continuous distractor free recall. In particular, in line with many dual component models of memory it suggests that different patterns of results should occur when examining retrieval from primary memory versus retrieval from secondary memory. Generally, evidence for or against dual component models has been gleaned from examining how different variables affect different aspects of the serial position curve. For instance, in immediate free recall, the recency effect has been taken as evidence for unloading from primary memory, while the lower probability of recalling items from the prerecency portion of the curve has been taken as evidence of retrieval from secondary memory (e.g., Atkinson & Shiffrin, 1968). The absence of a recency effect in delayed free recall has been taken as evidence that a distracting task selectively disrupts retrieval from primary memory while sparing retrieval from secondary memory (e.g., Glanzer, 1972). However, evidence that recency effects could be found in continuous distractor free recall argued against differences between primary and secondary memory (e.g., Bjork & Whitten). This finding has lead many to argue for a unitary memory system rather than a dual component memory system (e.g., Crowder, 1982; Nairne, 2002). Recently, however work by Davelaar and colleagues (Davelaar et al., 2005; Davelaar, Haarmann, Goshen-Gottstein, & Usher, in press) has suggested that recency effects in immediate and continuous distractor free recall are due to separate mechanisms. Similar to traditional dual component models, Davelaar et al. (2005) have argued that recency effects found with immediate free recall are due to unloading from a short-term buffer of some sort. Recency effects found in continuous distractor free recall, however, are due to differences in temporal-contextual retrieval from secondary memory. In support of this, Davelaar et al. showed that PI affected recency in the continuous distractor task but not in the immediate free recall tasks. Consistent with previous work (e.g., Craik & Birtwistle, 1971) Davelaar et al. argued that items within the short-term buffer were protected from PI and only items that had to be retrieved from secondary memory would be affected by PI. Thus, PI would affect the prerecency portion of the immediate free recall curve, but would affect all positions in the continuous distractor free recall curve.

According to the view presented in the present paper, this suggests that the serial position curves for high and low spans should change as a function of task. Specifically, because low spans are more susceptible to PI than high spans (due to noisier temporal-contextual cues) low spans should show impaired performance compared to high spans in those positions thought to reflect retrieval from secondary memory. For positions thought to reflect retrieval from primary memory, however, small to nonexistent span differences should occur. Additionally, as with previous research there should be a strong recency effect in immediate free recall, no recency effect in delayed free recall, and an attenuated recency effect in continuous distractor free recall. Figure 8 shows serial position functions for immediate (from Unsworth & Engle, 2005d), delayed, and continuous distractor free recall as a function of span for a list-length of twelve.

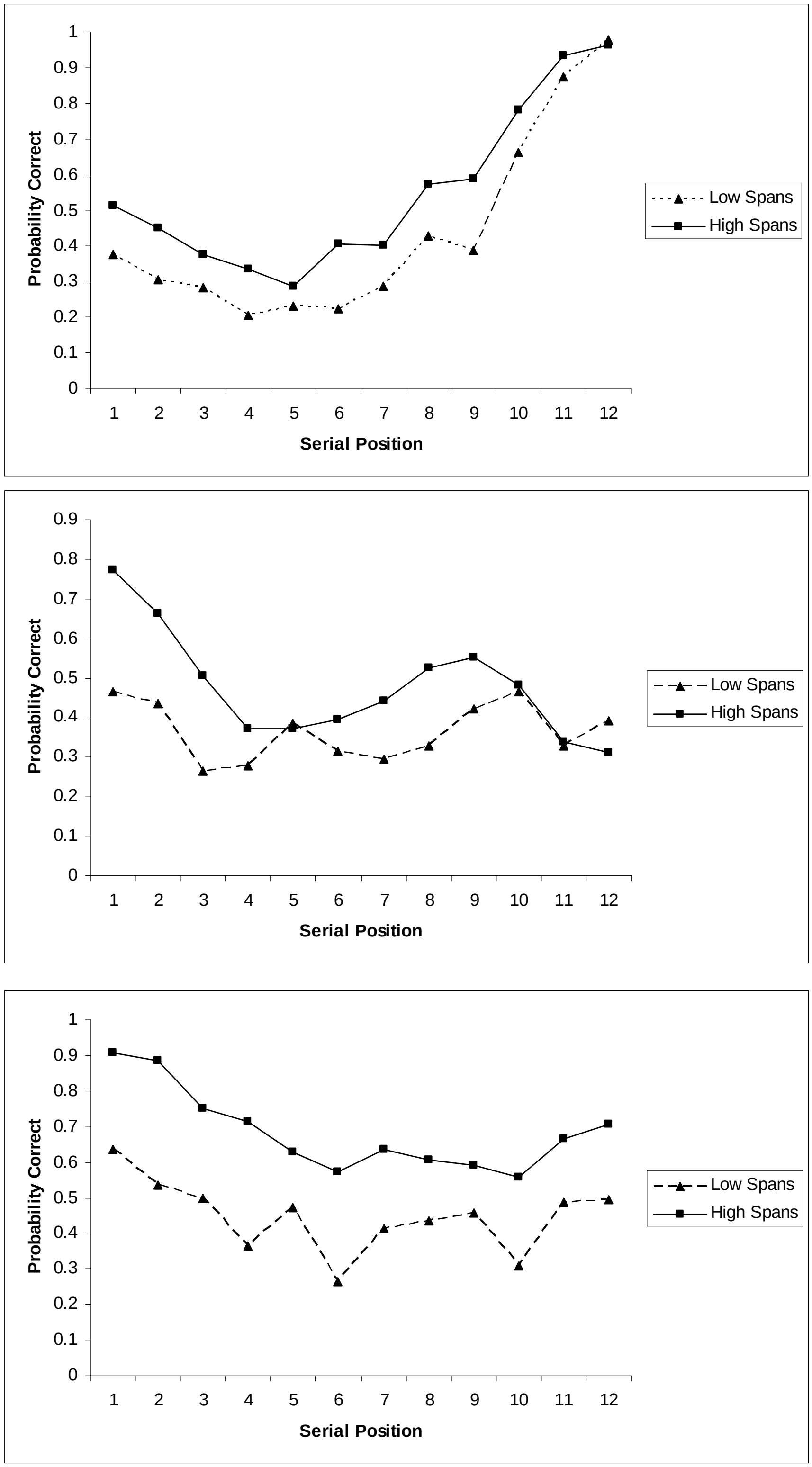

Figure 8. Probability correct as a function of serial position and complex span for immediate, delayed, and continuous distractor free recall. Immediate free recall adapted from Unsworth and Engle (2005d).

Consistent with the current view, highs and lows show small differences at recency positions, but larger differences for prerecency positions in immediate free recall. For delayed free recall, highs and lows differ at nearly all positions and the recency effect is abolished. For continuous distractor free recall the recency effect begins to reemerge and span differences are evident at all serial positions. These results are consistent with the view that immediate free recall differs from delayed and continuous distractor free recall in that some items in immediate free recall can be retrieved from primary memory, whereas in delayed and continuous distractor free recall all items have to be recalled from secondary memory. These results are, also, consistent with the view that low spans are poorer than high spans at retrieving items from secondary memory due to an increased search set size.

Additional evidence for this view comes from an examination of IRTs. According to the current view, if the first few items in immediate free recall are unloaded from primary memory, then these items should be associated with relatively rapid and equivalent IRTs. Furthermore, assuming that high and low spans do not differ in the time to unload items from primary memory, there should be no span differences in the first few IRTs. In support of these notions, Unsworth and Engle (2005d) found that the first few items recalled in immediate free recall were associated with rapid and nearly equivalent IRTs. Furthermore, this was true for both high and low span individuals. However, beginning with the third IRT, span differences emerged and IRTs became progressively slower with each successive item recalled. Unsworth and Engle argued that the first few items were unloaded from primary memory, while subsequent items were recalled from secondary memory, leading to increases in IRTs and span differences.

If this is the case, then an examination of IRTs in delayed and continuous distractor free recall, where no items are theoretically recalled from primary memory, should demonstrate the same pattern of IRTs as immediate free recall after getting rid of the first few IRTs. That is, those IRTs associated with secondary memory retrieval in immediate free recall should show the same pattern as IRTs for all items in delayed and continuous distractor free recall. Figure 9 shows IRTs growth functions for immediate (from Unsworth & Engle, 2005d), delayed, and continuous distractor free recall. Note that the IRTs for immediate free recall include only those IRTs associated with recall from secondary memory (i.e., the first two IRTs have been omitted). The results suggest that the IRT growth functions are similar for the three tasks when only considering IRTs associated with items retrieved from secondary memory. In all cases, IRTs get progressively slower and span differences are apparent. As with the serial position functions, these results are consistent with a dual component model of memory and with the notion that low spans search through a larger set of items in secondary memory than high spans leading to low spans' recall deficits.

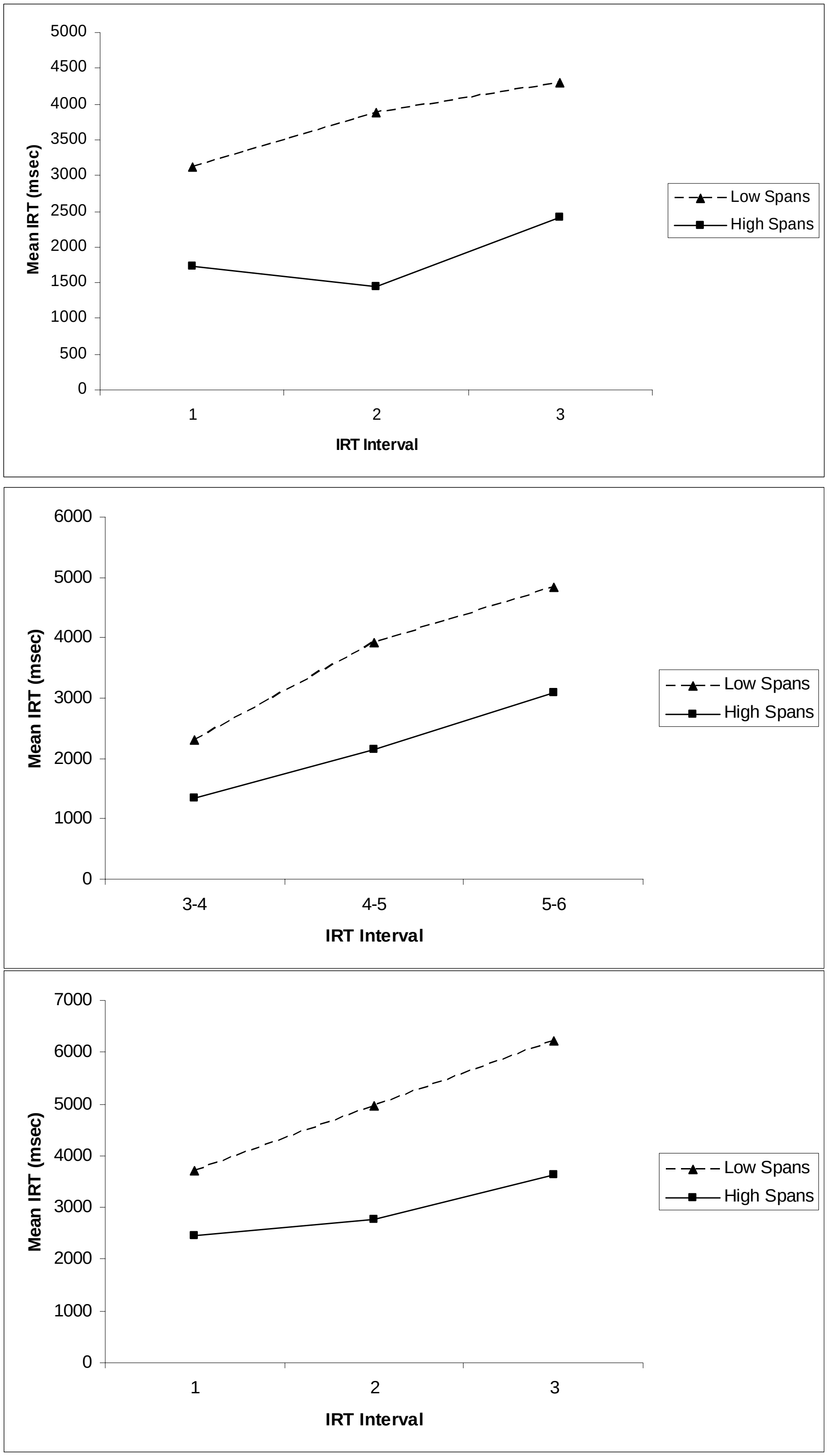

Figure 9. IRTs for three IRT intervals as a function of complex span for immediate, delayed, and continuous distractor free recall. Immediate free recall adapted from Unsworth and Engle (2005d).

Conclusions

Those individuals who score low on complex span measures of WMC have impaired retrieval from episodic memory compared to individuals who score high on these measures. In two experiments this deficit was shown to be related not only to the recall of fewer items, but also to the greater recall of previous list items from the immediately preceding list. The retrieval deficit was, also, associated with slower recall of items throughout the recall period, with low span individuals consistently recalling items at a slower rate than high span individuals. Collectively, these results support the notion that low span individuals are poorer at using temporal-contextual cues to focus the search of secondary memory on current target items. This inability to adequately focus the search results in many irrelevant items being included in the search set and leads to subsequent lower probabilities of recall and slower overall recall. Low spans retrieval deficits, therefore, are linked to deficits in using cues to guide the search process of secondary memory.

APPENDIX

Table A1. Mean total correct for each task by complex span for Experiment 1.

	Task			
Complex Span	Ospan	Symspan	Rspan	Raven
High Span	67.30 (5.58)	34.78 (3.61)	68.43 (3.69)	9.22 (1.62)
Low Span	37.15 (12.59)	19.65 (6.60)	30.60 (13.59)	6.85 (2.89)

Note. Ospan = operation span; Symspan = symmetry span; Rspan = reading span; Raven = Raven Advanced progressive matrices. Ospan and Rspan are out of 75, Symspan is out of 42, and Raven is out of 12. Numbers in parentheses are standard deviation.

Table A2. Frequency of participants enrolled in college by complex span for Experiment 1.

	College		
Complex Span	Georgia Tech	Other	None
High Span	15	4	4
Low Span	7	5	8

Note. Georgia Tech = Georgia Institute of Technology; Other = other Atlanta area colleges; None = not currently enrolled in college.

Table A3. Mean total correct for each task by complex span for Experiment 2.

	Task			
Complex Span	Ospan	Symspan	Rspan	Raven
High Span	66.75 (6.17)	35.55 (2.70)	67.30 (3.63)	10.20 (1.28)
Low Span	38.15 (13.52)	17.35 (6.16)	31.55 (12.93)	7.00 (2.66)

Note. Ospan = operation span; Symspan = symmetry span; Rspan = reading span; Raven = Raven Advanced progressive matrices. Ospan and Rspan are out of 75, Symspan is out of 42, and Raven is out of 12. Numbers in parentheses are standard deviation.

Table A4. Frequency of participants enrolled in college by complex span for Experiment 2.

	College		
Complex Span	Georgia Tech	Other	None
High Span	12	3	5
Low Span	4	7	9

Note. Georgia Tech = Georgia Institute of Technology; Other = other Atlanta area colleges; None = not currently enrolled in college.

REFERENCES

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

Anderson, M.C., & Spellman, B.A. (1995). On the status of inhibitory mechanisms in Cognition: Memory retrieval as a model case. *Psychological Review, 107,* 68-100.

Atkinson, R. C. and Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation, Vol. II* (pp.89-195). New York: Academic Press.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133,* 83-100.

Barrouillet, P., & Camos, V. (2001). Developmental increase in working memory span: Resource sharing or temporal decay? *Journal of Memory and Language, 45,* 1-20.

Bjork, R.A., & Whitten, W.B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173 – 189.

Bousfield, W.A., Sedgewick, C.H.W., & Cohen, B.H. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology, 67,* 111-118.

Burns, D.J., & Schoff, K.M. (1998). Slow and steady often ties the race: The effects of item-specific and relational processing on cumulative recall. *Journal of ExperimentalPsychology: Learning, Memory, and Cognition, 24,* 1041-1051.

Conway, A.R.A., & Engle, R.W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*, 354-373.

Conway, A.R.A., Cowan, N., & Bunting, M.F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review, 8*, 331-335.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B.

(2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30,* 163-183.

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

Cowan, N. (2005). Working memory capacity limits in a theoretical context. In C. Izawa & N. Ohta (Eds.), *Human learning and memory: Advances in theory and application. The 4th Tsukuba international conference on memory.* Erlbaum.

Cowan, N., Johnson, T.D., & Saults, J.S. (2005). Capacity limits in list item recognition: Evidence from proactive interference. *Memory, 13,* 293-299.

Craik, F.I.M. & Birtwistle, J. (1971). Proactive inhibition in free recall. J*ournal of Experimental Psychology, 91,* 120–123.

Crowder, R.G. (1982). The demise of short-term memory. *Acta Psychologica, 50,* 291-323.

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

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25,* 1 – 18.

Davelaar, E.J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H.J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112,* 3-42.

Davelaar, E.J., Haarmann, H.J., Goshen-Gottstein, Y., & Usher, M. (in press). Semantic similarity dissociates short- from long-term recency effects: Testing a neurocomputational model of list memory. *Memory & Cognition.*

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,* 309-331.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation, Vol 5.* New York: Academic Press.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351-60.

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,* 3-15.

Halford, G.S, Mayberry, M.T., & Bain, J.D. (1988). Set-size effects in primary memory. An age related capacity limitation? *Memory & Cognition, 16,* 480–487.

Hasher, L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp.193 – 225). San Diego, CA: Academic Press.

Hasher, L., Zacks, R.T., & May, C.P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention & Performance, XVII, Cognitive Regulation of Performance: Interaction of Theory and Application* (pp. 653-675). Cambridge, MA: MIT Press.

Heitz, R.P., & Engle, R.W. (2004). *Focusing the spotlight: Individual differences in visual attention control.* Manuscript submitted for publication.

Herrmann, D.J., & Chaffin, R.J.S. (1976). Number of available associations and rate of association for categories in semantic memory. *Journal of General Psychology, 95,* 227-231.

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

Indow, T., & Togano, K. (1970). On retrieving sequences from long-term memory. *Psychological Review, 77,* 317-331.

Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130,* 169-183.

Kane, M.J., Conway, A.R.A., Hambrick, D.Z., & Engle, R.W. (in press). Variation in working-memory capacity as variation in executive attention and control. To appear in A.R.A. Conway, C. Jarrold, M.J. Kane, A. Miyake, & J.N. Towse (Eds.), *Variation in working memory.* New York: Oxford University Press.

Kane, M.J., & Engle R.W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 333 – 358.

Kane, M.J., & Engle, R.W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47 – 70.

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. *Journal of Experimental Psychology: General, 133,* 189-217.

Kaplan, I.T., Carvellas, T., & Metlay, W. (1969). Searching for words in letter sets of varying size. *Journal of Experimental Psychology, 82,* 377-380.

Klein, K., & Boals, A. (2001). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology 15*, 565-579.

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

Kyllonen, P.C., & Stephens, D.L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences*, 2, 129 – 160.

Long, D.L., & Prat, C.S. (2002). Working memory and Stroop interference: An individual differences investigation. *Memory & Cognition, 30,* 294-301.

May, C.P., Hasher, L., & Kane, M.J. (1999). The role of interference in memory span. *Memory & Cognition, 27,* 759 – 767.

McGill, W.J. (1963). Stochastic latency mechanism. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol 1., pp. 309-360). New York: Wiley.

Metlay, W., Handley, A., & Kaplan, I.T. (1971). Memory search through categories of varying size. *Journal of Experimental Psychology, 91,* 215-219.

Murdock, B. B. (1972). Short-term memory. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol 5. New York: Academic Press.

Nairne, J.S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53,* 53 – 81.

Naveh-Benjamin, M., & Guez, J. (2000). Effects of divided attention on encoding and retrieval processing: Assessment of attentional costs and a componential analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1461-1482.

Raaijmakers, J.G.W., & Shiffrin, R.M. (1980). SAM: A theory of probabilistic search of associative memory. In G. Bower (Ed.). *The psychology of learning and motivation, Vol 14.* New York: Academic Press.

Raaijmakers, J.G.W., & Shiffrin, R.M. (1981). Search of Associative Memory. *Psychological Review, 88,* 93-134.

Roediger, H.L. III, Stellon, C.C., & Tulving, E. (1977). Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning & Memory, 3,* 174-188.

Redick, T.S., & Engle, R.W. (in press). Working memory capacity and attention network test performance. *Applied Cognitive Psychology.*

Rohrer, D. (2002). The breadth of memory search. *Memory, 10,* 291-301.
Rohrer, D., & Wixted, J.T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition, 22,* 511 – 524.

Rosen, V.M. & Engle, R.W. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39*, 418 – 436.

Rosen, V. M., Bergeson, J.L., Putnam, K., Harwell, A., and Sunderland, T. (2002). Working memory and apolipoprotein E: What's the connection? *Neuropsycholgia, 40*, 2226-2233.

Schamader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85,* 440-452.

Shiffrin, R.M. (1970). Memory search. In D.A. Norman (Ed.), *Models of Human Memory* (pp. 375 – 447). New York: Academic Press.

Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127-154.

Unsworth, N., & Engle, R.W. (in press). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language.*

Unsworth, N., & Engle, R.W. (2005b). Working memory capacity and fluid abilities: Examining the correlation between operation span and raven. *Intelligence, 33,* 67-81.

Unsworth, N., & Engle, R.W. (2005b) Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition, 33,* 213-220.

Unsworth, N., & Engle, R.W. (2005c). *A temporal-contextual retrieval account of complex span: An analysis of errors.* Manuscript submitted for publication.

Unsworth, N., & Engle, R.W. (2005d). *The nature of individual differences in working memory capacity: Capacity of primary memory and retrieval from secondary memory.* Manuscript in preparation.

Unsworth, N., Heitz, R.P., & Engle, R.W. (2005). Working memory capacity in hot and cold cognition. In R.W. Engle, G. Sedek, U. Hecker, & D.N. McIntosh (Eds.) *Cognitive limitations in aging and psychopathology: Attention, working memory, and executive functions* (pp. 19-43). New York: Oxford University Press.

Unsworth, N., Heitz, R.P., Schrock, J.C., & Engle, R.W. (in press) An automated version of the operation span task. *Behavior Research Methods.*

Unsworth, N., Schrock, J.C., & Engle, R.W. (2004) Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1302-1321.

Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology, 31,* 1-23.

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

Watkins, M.J. (1979). Engrams as cuegrams and forgetting as cue over-load: A cueing approach to the structure of memory. In C.R. Puff (Ed.), *Memory organization and structure* (pp. 347-372). New York: Academic Press.

Wixted, J.T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1024 – 1039.

Wixted, J.T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An intergrative review of the empirical literature. *Psychonomic Bulletin & Review, 1,* 89-106.