The contribution to immediate serial recall of rehearsal, search speed, access to lexical memory, and phonological coding: An investigation at the construct level

Gerald Tehan, Gerard Fogarty, Katherine Ryan

University of Southern Queensland

Toowoomba

Australia

MS03-175R

Mailing Address:
Gerry Tehan, PhD.
Department of Psychology
University of Southern Queensland
Toowoomba, 4350
AUSTRALIA
Phone number: + 61 7 4631 2376
Fax: + 61 7 4631 2721
e-mail: tehan@usq.edu.au

Full reference: Tehan, G., & Fogarty, G. (2004). Individual Difference in Memory Span: The Contribution of Rehearsal, Access to Lexical Memory and Phonological Codes. *Memory and Cognition*, 32 (5), 711-721.

Abstract

Rehearsal speed has traditionally been seen to be the prime determinant of individual differences in memory span. Recent studies, in the main using young children as the participant population, have suggested other contributors to span performance, notably contributions from long-term memory and forgetting and retrieval processes occurring during recall. In the current research we used structural equation modelling to explore at the construct level individual differences in immediate serial recall with respect to rehearsal, search, and speed of access to lexical memory. We replicate standard short-term phenomena; we show that the variables that influence children's span performance influence adult performance in the same way; and we show that lexical memory access appears to be a more potent source of individual differences in immediate memory than either rehearsal speed or search factors.

The Contribution to Word Memory Span of Rehearsal, Search Speed, Access to Lexical Memory, and Phonological Coding: An Investigation at the Construct Level.

Most current models of immediate serial recall have three basic assumptions. Firstly, that short-term information is lost very rapidly and as a result serial recall is often based upon a degraded memory trace. Some models attribute forgetting to decay (Baddeley, 1986; Burgess & Hitch, 1999; Cowan, 1999; Page & Norris, 1999), others see interference as the cause of forgetting (Brown & Hulme, 1995; Nairne, 1990). The second assumption is that in order to recall a list item, a degraded memory trace must be reconstructed or redintegrated in some manner (Brown & Hulme, 1995; Farrell & Lewandowsky, 2002; Schweickert, 1993). Thirdly, the assumption is made that the trace that supports recall is speech-based in nature. That is, immediate recall relies upon phonological codes. While these models assert that these features are important determinants of immediate recall, they say little about the relative importance of each factor or what contribution to span each factor makes. One way of addressing this problem is to take an individual differences approach and to assess the contribution of each of these factors to immediate serial recall (ISR). This approach is adopted in the current study.

The standard account of immediate memory (Nairne, 2002) sees performance to be the outcome of a balancing act between rapid de-activation of the memory trace and rehearsal, where rehearsal is assumed to refresh and re-activate the fading trace. Many of the variables that are believed to influence memory span are thought to have their influence via the ease with which rehearsal can be carried out and the memory trace refreshed.

The concept of rehearsal has been notoriously difficult to measure and most researchers have opted for indirect measures, the most widely adopted being speed of overt articulation. In probably the first study to use this indirect measure of rehearsal, Baddeley, Thomson and Buchanan (1975) had their participants repeat three words as quickly as possible ten times and, as a second measure, read a list of 50 words as quickly as possible. Using regression procedures they found a strong relationship between immediate memory span and the speeded articulation measures; span could be predicted on the basis of the number of items a participant could articulate in approximately 2 seconds. The correspondence between articulation and span was taken as prima facie evidence that rehearsal speed was a prime determinant of span. Span-rehearsal rate correspondences have subsequently been used to account for many of the structural (Schweickert & Boruff, 1986; Tehan & Humphreys, 1988), developmental (Hulme, Thomson, Muir & Lawrence, 1984), cross-cultural (Chen & Stevenson, 1988), and individual differences that are observed in span performance (Baddeley et al. 1975).

Despite the initial appeal of the rehearsal speed explanation, recent evidence suggests that its role is not as dominant as first thought. For a start, there are many instances where rehearsal rate and span dissociate. It is possible to match items for spoken duration and still see differences in span (Hulme, Maughan & Brown, 1991).

Likewise, it is possible to see differences in spoken duration without accompanying changes in span (Caplan, Rochon & Waters, 1992). In the light of these and similar findings, it has been suggested that rehearsal may not actually be all that important (Brown & Hulme, 1995; Nairne 2002).

Cowan and his colleagues (Cowan, Day, Saults, Keller, Johnson, & Flores, 1992) presented an alternative account of the rapid forgetting in immediate memory that is based upon decay during the recall process. They demonstrated that longer words at the start of the list caused more forgetting for the words at the end of a list than if short words were presented at the start of the list. Cowan argued that the trace that supported recall decayed during the recall process and that more decay occurred during the articulation of long words during output than short words. From this perspective, articulation speed was an indicator not of rehearsal speed but of the amount of rapid decay that would occur during output. If a person can rapidly articulate the remembered item, decay of the non-recalled items would be minimised.

Furthermore, through their exploration of the timing characteristics of verbal output of subspan and span length lists, Cowan et al. (1994) noted that as lists increase in difficulty, participants produce longer and longer pauses between list items. Given a memory architecture where a limited number of items are rapidly returning to threshold levels of activation, they proposed that during pauses participants reactivated the memory trace and searched through these activated traces for the next response. Cowan, Wood, Wood, et al. (1998) tested these ideas using an individual differences approach and were able to show via structural equation modelling (SEM) that articulation and search measures both made significant and independent contributions to span. In a similar study, Hulme, Newton, Cowan, Stuart and Brown (1999) also explored the relationship between span and memory search rate. They found a significant correlation between search rate and span for one syllable words, but this relationship did not hold for 5 syllable words, nor for one and five syllable non-words. As such, the role of memory search rate is worthy of further exploration.

It is now commonly accepted that a critical process in immediate serial recall is the ability to reconstruct or redintegrate a degraded short-term trace. While current models of immediate memory incorporate such a stage (Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Page & Norris, 1999), these models are not overly specific about the memory system that is accessed during this process. It is becoming increasingly clear that redintegration involves accessing lexical and/or semantic memory (Hulme et al., 1991; Schweickert, 1993). Lexical features of verbal materials often have a similar impact upon immediate recall. Most obviously, words are better recalled than non-words (Hulme et al., 1991). Word-frequency likewise has an impact upon span performance that is not mediated by rehearsal (Tehan & Humphreys, 1988). Content words (e.g., nouns, verbs, and adjectives) are recalled better than function words (e.g., prepositions, pronouns, articles) under both rehearsal and suppression conditions (Tehan & Humphreys, 1988). Results such as these have led to the proposition that at least two components underlie span performance, a rapidly decaying phonemic trace which is augmented by long-term knowledge about the lexical attributes of words (Schweickert, 1993; Brown & Hulme, 1995; Hulme et al., 1991).

The reason for assuming that lexical access involves phonemic traces is based on the finding that immediate recall seems to be based primarily upon phonological representations. It has long been known that participants have difficulty remembering lists of items that rhyme or have substantial phonological overlap (Baddeley, 1966). Given the ubiquitous role of phonological codes in immediate recall, Hulme et al. assumed that long-term phonological information would be accessed to facilitate the reconstruction of degraded phonological representations. Moreover, at least in the developmental literature, there is substantial evidence that children’s ability to process phonological information is a substantial contributor to span (Kail, 1997).

In summary, the notion that rehearsal is the sole or even major contributor to span has to be questioned. Decay during recall, speed of search, the speed of access to lexical memory, and facility with phonological coding have all been proposed as alternative or additional mechanisms. Tehan and Lalor (2000) used an individual differences approach to explore the relationships between some of these measures, namely rehearsal, speed of lexical access, and output time. With multiple measures of each construct, a factor analysis supported the three latent constructs. Tehan and Lalor attempted to gauge the relative contributions of the three factors to memory span by conducting a series of hierarchical regression analyses and varying the order of entry of the factor scores. In their first study, they found that all three factors made significant contributions to memory span but that Lexical Access contributed most to digit span scores when it was the first factor entered in the equation (29%). Rehearsal and output measures contributed equivalent amounts to span (15%) when they were each entered first in the regression equation. In a second study, they found that Lexical Access accounted for 19% and Rehearsal for a mere 4% of the variance in span.

Although they were able to produce strong evidence for the importance of speed of lexical access in memory span, the Tehan and Lalor study was confined to the comparison of speed of lexical access, rehearsal, and output speed (Study 1), and to lexical access and rehearsal (Study 2). No measures of search speed or facility with phonological codes were included. As pointed out earlier in the introduction to this paper, both of these variables have been linked with memory span. Cowan et al. (1998) argued for the importance of search processes and Hulme et al. (1991) emphasised the role of phonological coding.

While some of these variables have been considered in combination, no single study has combined all four factors. In the following study, the rehearsal and speed of lexical access measures used by Tehan and Lalor were included along with measures of speed of search, and measures of phonological coding in an attempt to determine the relative contribution of these factors to immediate recall.

Issues with Measures

The above aims can be adequately evaluated only if valid markers of the variables are selected. We have indicated what some of those markers might be but it is important to justify our selection because, in many instances, prior research has raised significant doubts about some of the measures used.

Rehearsal: In the original work exploring the relationship between span and rehearsal rate (Baddeley et al., 1975), two indirect measures of rehearsal were utilised: reading a list of 50 words as quickly as possible and repeating three words as quickly as possible ten times. The reading time estimate of rehearsal speed has faded from use due to the complexity of the reading task. Moreover, in the Tehan and Lalor (2000) data the two tasks loaded upon different factors. So the two tasks are not equivalent.

The repetition task has been widely adopted as a measure of maximal articulation rate/rehearsal speed. This measure is not free of problems either. For example, Chase, Lyon and Ericsson (1981) reported two studies in which rehearsal of 3, 4 or 5 digit lists did not correlate with span whereas rehearsing a list of six digits did correlate with span. Because the correlation emerged only as span was approached, they argued that the correlation of span with articulation rate is an artefact of the involvement of memory for order in both tasks. Hulme et al. (1984) explored these issues by measuring articulation rates for single words and groups of three words in a sample of children. They found that both measures correlated equally well with memory performance. Likewise, Ferguson, Bowey & Tilley (2002) explored speech rate in children for one word and three word lists. Like Hulme et al, they found that both measures correlated with span with the correlation higher for the three word lists. Using regression techniques they found that if speech rate for single item words was partialled out of speech rate for the word triples, the residuals were still correlated with span. They concluded, as did Chase et al., that the speeded articulation of word triples involved a memory component that inflated the correlation between span and articulation rate.

The above research suggests that one needs to be careful in the selection of speeded articulation measures so as to ensure that they have minimal memory load. Consequently, we have adopted those tasks used by Lalor and Tehan (2002). They used the triple word repetition task but supplemented these with two tasks that involved the speeded articulation of overlearned sequences that should not involve a memory load component. Thus participants were required to recite the letters of the alphabet as quickly as possible and to count backwards from 20 to 1 as quickly as possible. These latter two measures are assumed to be free of a memory load component.

Lexical Access: One of the key assumptions of the widely proposed redintegration hypothesis is that lexical memory is addressed with the aim of reconstructing degraded phonological information. The role of speeded access to lexical memory in visual word recognition has been studied using four principal tasks: Lexical decision, speeded word naming, speeded nonword naming, and perceptual identification. While some might wonder about the inclusion of non-word naming as a measure of lexical access, given that by definition non-words have no lexical representations, nonword naming has been important to the development of “dual route” models of lexical access (Coltheart, 1980).

Tehan and Lalor (2002) examined the relationship between the first three tasks and memory span. They demonstrated that the three tasks loaded on the same factor in their battery of tasks and that all three measures were strongly related to span. In fact, in their data, these measures continually produced a robust contribution to individual differences in span performance. Consequently, we have adopted these tasks in the current study.

In adopting these speeded tasks, it is clear that it is speed of access to lexical memory that we see as the important aspect of the redintegration process. Other aspects of lexical access like the extent of the development of lexical memory, age of acquisition or neighbourhood effects may have an effect upon span, but are not the focus of the current research.

Phonological Coding: The role of individual differences in phonological coding has most often been studied in the context of children’s ability to read. The individual differences studies often explore the relationship between phonological coding ability, memory span and some measure of reading ability. In these studies there are quite a range of measures of phonological ability measures. Some of the more common tasks involve phoneme deletion (“what word results if the first sound of the word spark is deleted?”); non-word repetition (say the word “bliborg”), phoneme blending where the ability to blend a sequence of sounds into non-words is measured; non-word segmentation, where children are asked to identify the component phonemes in a nonword; sound categorisation which measures the child’s ability to detect rhyme and/or alliteration (which of the following words does not fit “cat, hat, man, fat”.

While these tasks all deal with phonological aspects of words, not tasks necessarily tap the same construct. For example, Hatcher and Hulme (1999) found in a factor analysis of many of the above tasks, that phoneme blending, phoneme segmentation, phoneme deletion and non-word repetition loaded on a separate factor to the sound categorisation task.

Many of the above tasks seemed to be somewhat unsuitable for use with adults as participants. To this end we consulted the adult individual differences literature to see if there were similar tasks to those used in the child literature. Carroll (1993) conducted a meta-analysis of the major factorial studies of human abilities and identified a number of tasks that measured phonological coding. Two of the tasks that loaded on the phonological coding factor were the Sound Grouping Test (Thurstone, 1936) and the Turse Phonetic Association test (Turse, 1940). Variants of these tests were constructed for the current study.

The sound grouping test is almost identical to the sound categorisation task that is commonly used in the developmental literature. Participants are presented with four words that look very similar to each other but only three words rhyme (pork, work, fork, cork). The task is to select the non-rhyming word. Given the similarity between the sound grouping task for adults and the sound categorisation task for children, we were reasonably confident that we were measuring the same construct. Given the ubiquitous finding in the developmental literature that the sound categorisation task correlated with span, we expected that at the very least the sound grouping task would also correlate with span.

The phonetic association test is basically a pseudohomophone naming task in which participants are presented with a non-word that if pronounced according to grapheme-phoneme conversion rules results in a pronounceable English word (e.g. kayj for cage). Again, this task has face validity as measure of phonological processing in that participants must convert phonological information into lexical information.

Search: As indicated earlier, Cowan et al. (1998) found that span was related to memory search rate in a Sternberg type task where search rate was determined by the slope of the reaction time set size function. While the search assumptions of the task are not theoretically neutral, our concerns are with the procedural aspects of the task. If the rehearsal measures can be criticised on the basis of contamination via a memory component, then this argument applies equally well to the Sternberg measure in general (Hulme et al., 1999) and the task that Cowan et al, (1998) employed. In both the Hulme and Cowan studies there are times where participants are holding four or five items in memory. For example, in the Cowan et al. task, participants were presented with a matrix of digits or letters. They were then given either 1, 3 or five items as probes and had to quickly search through the matrix and check all instances of the probes. On the basis of the reaction times to the different number of probes Cowan et al., were able to calculate search rates.

The crucial feature of this task for present concerns is that participants must maintain, on some trials, up to four or five items in memory. Thus, it is clear that both span and search tasks involve a substantial storage component and it may well be this feature that is producing the correlation.

If one allows the possibility that the correlation between span performance and performance on the Sternberg task is due to common storage factors, then it becomes difficult to provide a more direct measure of memory search where memory is not involved. As is the case with rehearsal, one is left with indirect measures of search speed. One way that this might be done is to present a list of items on a page, present a probe and time how quickly participants can find this probe. This is the approach we adopt in the current study. That is, we use a simple search task, one that is not overly contaminated by memory storage. In the developmental literature, search tasks of this type (e.g. the cross-out task, and identical pictures test) are correlated with span performance (Kail, 1997), so there is some precedent for expecting that simple search measure may well contribute to span performance. However, the possibility remains that the search involved in our task has nothing to do with the search processes that take place in memory.

Immediate Memory: The most commonly used measure of short-term memory is digit span. However, digits are highly familiar with very limited semantic representations (Hulme & Roodenrys, 1995). Clinical studies have shown that children with phonological difficulties can achieve essentially normal performances on digit span tasks, but show deficits on word span tasks (Snowling & Hulme, 1989). These findings have led some researchers to conclude that word span is probably a more sensitive test than digit span (Hulme & Roodenrys, 1995). Consequently, in this study participants were required to recall short lists of words. Furthermore, we use a large pool of words as the stimulus material, such that each word appears in the memory tasks only once. This is in contrast to the closed set typically used in digit or letter span.

We also see that there are two further drawbacks to using the span measure. The span measure very often represents the longest list that has been recalled without error. Given that errors can be informative of underlying processes, the use of a span measure only provides one source of information. More importantly, however, most current models of immediate recall are not models of span. They are models of list recall where errors occur. That is, they are concerned not only with recalling items in their correct position, but are equally focused upon explaining the various patterns of errors that are found in immediate serial recall. These models are thus implicitly based upon trials of a fixed length that is slightly above span for most people. The correlates of fixed list performance will obviously have direct implications for these models. We are assuming here that span scores and correct in position scores of fixed lists are highly correlated and share common processes. However, it is important to confirm that this is actually the case.

Research Design and Hypotheses

Whereas Tehan and Lalor used exploratory factor analysis followed by hierarchical regression analysis in both the studies they reported, structural equation modelling (SEM) was used in the present study. It was expected that the covariances among the 12 marker variables could be explained by five underlying factors, as specified by the measurement model. The structural part of the model showed the four correlated factors predicting the Memory Span latent construct. Following the work of Tehan and Lalor, it was expected that lexical access and rehearsal factors would make significant independent contributions to the prediction of memory span. Based on the literature, additional contributions were expected from search and phonological coding processes.

Method

Participants

There were 126 participants, comprising students who participated for course credit and community volunteers who received a ticket in a small cash lottery.

Measures

The three speeded articulation measures used by Tehan and Lalor (2000) were used here as well. The three tasks are the repetition of three-words, alphabet recitation and counting time. Single trials were used on all three rehearsal tasks.

1. Articulation Time. Participants were required to repeat the words "Blanket Summer Friday" as quickly as possible for 10 repetitions. A stopwatch was used to measure the number of seconds taken for the 10 repetitions.

2. Counting Time. Participants were instructed to count backwards from 20 to 1 as fast as they could. A stopwatch was used to measures the number of seconds taken for the task.

3. Alphabet Recitation. This task involved speeded verbalisation of the alphabet from A-Z. Again, a stopwatch was used to record the number of seconds required for this task.

The lexical access measures were those used by Tehan and Lalor (2000) and included a lexical decision task, a non-word naming task, and a word naming task.

4. Lexical Decision. Participants were presented with a list of 40 low-frequency, 5-letter strings from the Toronto Word Pool norms (Friendly, Franklin, Hoffman, & Rubin, 1982). The letters in 20 of these words were substituted to create phonotacticly legal nonwords (e.g., abort → amort). The newly created nonwords were then randomly interspersed among the remaining words. The list was divided into 4 columns and displayed on the computer screen in its entirety for 15 seconds. Participants were required to look at each letter string and to decide as quickly as possible whether the letter string was an English word or not. Lexical decision performance was calculated by measuring the number of correct decisions within the 15 second time limit.

5. Nonword Naming. The 42 stimulus items for this task were constructed in the same way as the nonwords used in the lexical decision task. Participants were presented with 3 columns of 14 nonwords and instructed to read aloud down each column as quickly as possible, pronouncing each non-word aloud. The nonword list disappeared from the computer screen after 15 seconds. Nonword naming scores were calculated by summing the number of nonwords pronounced within the 15 second time limit.

6. Word Naming. The materials for the task included 50 low-frequency, 5-letter words from the Toronto Word Pool norms. The word list was divided into 5 columns and participants were required to read the words aloud moving down the columns until the list disappeared from the computer screen. A word naming score was calculated by summing the total number of words read correctly within the 15 second time interval.

Carroll (1993) provided the tasks used to assess phonological coding. Two of these were the Sound Grouping Test (Thurstone, 1936) and the Turse Phonetic Association test (Turse, 1940). Variants of these tests were constructed for the current study.

7. Sound Grouping. The Sound Grouping (Thurstone, 1936) test involves presenting participants with sets of visually similar words. Three of these words rhyme with one another but the fourth does not (e.g., pork-cork-work-fork). The version of the Sound Grouping used here consisted of 30 such lists. We used the lexical access literature to find a list of thirty regularly pronounced and 30 irregularly pronounced visually similar word pairs. We then used the University of South Florida rhyme norms (Walling, McEvoy, Oth, & Nelson, 1994) to find two other regularly pronounced rhymes for each of the selected pairs. Having established the word pool we randomly ordered the four words on each list and then randomised the sequence of 30 lists. All participants studied the same list. Participants were asked to quickly circle the word in each row of the sheet provided which sounded different to the other three items in the row. The dependent variable was the average number of seconds taken to select the irregular word from the three distractors.

8. Pseudo Words. The Turse Phonetic Association test (Turse, 1940) is basically a pseudohomophone naming task (nonwords which sound like words when pronounced e.g., phocks, kayj, durt, etc). We again accessed the lexical access literature to find appropriate examples of such pseudohomophones. A total of 42 pseudohomophones were presented in 3 columns on the computer screen. Participants were instructed that they would be presented with a series of non-words that when pronounced sounded like legitimate English words. Their task was to read aloud down each column as quickly as possible. Scores on the Pseudo Words task were calculated by summing the number of pseudohomophones correctly pronounced within the 15 second time limit.

Cowan et al., (1998) argued that search strategies were an important determinant of span. In the present study we used a word search task and a digit search task that involved search but minimal memory components.

9. Word search. The word search task was based on the Hidden Words task from Ekstrom, French, Harman, and Dermen (1976). Participants searched for English words that were embedded in a 15 by 15 letter matrix. The letters of a word were presented either forwards or backwards within a single row, a single column, or a diagonal. The list of target words was presented underneath the matrix. Participants were instructed to locate and circle the hidden words in the matrix. They were instructed to work as quickly as possible. The dependent variable was the number of words detected in a 20 second period.

10. Digit Search. This task consisted of an 11 by 12 matrix of the digits 1 to 20. Participants were instructed to go through the matrix and circle each instance of the number "4" and each instance of the number "13". Instructions stressed the speeded nature of the task and the dependent measure was the number of detections across a 20 second period.

*11. Immediate Serial Recall (ISR).*The task was initially set up to examine proactive interference effects using a methodology developed by Tehan and Humphreys (1995). In the current version of the task, participants studied lists that contained either one or two blocks of five words. They were instructed to remember only the most recent block of five words, but that they would not know in advance whether it would be a one-block trial or a two-block trial. Each trial started with an audible "beep" to signal that the words were about to be presented. The words were presented individually on a computer screen, at a rate of one word per second. On one-block trials the five words were presented and recall was immediately tested. In a two-block trial, an exclamation mark (!) was presented for one second after the fifth word in the first block. This was the signal that the trial was a two block trial and that participants were to forget the first block and concentrate on remembering the second block. A row of question marks "???" appeared on the screen directly following the final word in each list to indicate that the 7 second recall period had commenced. Participants were instructed to verbally recall the words in the correct order. The number of items correctly recalled in position on each list served as the chosen measure of serial recall performance.

Participants studied 20 trials, comprising an equal number of one-block and two-block trials, randomly interspersed. The study words consisted of low frequency, one syllable words from the University of South Florida rhyme category norms (Walling, McEvoy, Oth, & Nelson, 1994). Phonological similarity was manipulated by having half the one-block and half the two-block trials consist of items that all rhymed (e.g., bean-spleen-green-scene-mean). The remaining lists were made up of phonologically dissimilar words (e.g., pale-den-pink-ball-milk). The twenty trials were randomly ordered and this order was constant across participants.

Initial examination of the data tended to show that participants paid more attention to the second block than the first block in that there were recall differences between one-block and two-block trials with an advantage for the two-block trials. Although proactive interference effects were included in the design we decided not to explore these issues in this paper (primarily because there was very little evidence for PI). Instead we concentrated on the phonological aspects of performance. As such, the two dependent measures used were performance on the phonologically dissimilar two-block trials and performance on the two-block phonologically similar trials. The number of items recalled in their correct serial position was used as the dependent score.

Procedure

Testing was conducted on an individual basis under standard laboratory conditions. Sound Grouping was presented first followed by the visual search tasks. The remaining computer-based tasks were then administered in the following order: Word Span, Counting, Alphabet, Articulation, Lexical Decision, Word Naming, Nonword Naming, and Pseudo Words. Timing measures were obtained using a standard digital stopwatch and all verbal responses were recorded by the experimenter. Total testing time was approximately 30 minutes.

Results and Discussion

Prior to the main analyses, data were screened for out of range values, multivariate outliers, and both univariate and multivariate normality. Three cases were deleted as a result of this screening, leaving 123 cases in the data set. The four variables that used time measures as the dependent variable exhibited severe skewness and kurtosis, so log transformations were applied to all four. Following these data screening procedures, all variables, individually and collectively, met the assumptions for parametric analysis. Descriptive statistics and Pearson Product Moment correlations are shown in Table 1.

Table 1

Descriptive Statistics and Correlations for All Variables in Study 1 (N = 123)

	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Articulation	2.49	.16											
2. Counting	1.98	.25	.44										
3. Alphabet	1.62	.24	.48	.60									
4. Lexical	15.27	4.59	-.23	-.27	-.08								
5. N-W Name	17.43	5.80	-.36	-.37	-.20	.45							
6. Word Name	29.56	5.87	-.39	-.40	-.27	.47	.65						
7. Sound Group	4.88	.34	.35	.42	.24	-.45	-.48	-.52					
8. Pseudo Wds.	8.09	4.44	-.31	-.32	-.17	.39	.61	.54	-.55				
9. Word Search	2.76	1.54	-.06	-.10	.05	.07	.06	.07	-.11	.07			
10. Digit Search	8.76	2.23	-.07	-.30	-.17	.15	.11	.16	-.06	.05	.22		
11. Span Sim.	12.28	5.27	-.24	-.23	-.10	.38	.34	.30	-.41	.30	.16	.12	
12. Span Diss.	13.96	6.08	-.25	-.25	-.15	.35	.34	.33	-.44	.34	.08	.16	.74

Note: Correlations above 0.17 were significant at the .05 level.

Preliminary Analyses: Bench Mark Results.

The data shown in Table 1 can be used to compare findings in the current study with those reported by earlier researchers. Looking at the bottom two rows in Table 1, it can be seen that phonologically dissimilar words were better recalled than the phonologically similar words, t (122) = 4.46, p. <.05, thus replicating the phonological similarity effect. It is also clear from the bottom two rows in Table 1 that the pattern of correlations was very similar for recall of both phonologically dissimilar and similar lists. More importantly, the current findings replicate the rehearsal rate and lexical access relationships with span that have been observed in other data sets. For example, the present findings replicate the Baddeley et al. (1975) results in showing a significant correlation between serial recall and articulation rate and they support Tehan and Lalor (2000) in showing that all three lexical access measures correlated with recall. They also support the Tehan and Lalor (2000) finding that the time to recite the letters of the alphabet is not correlated with recall and show that articulation and speed of lexical access effects generalise from span measures to serial recall of fixed length lists. It is clear that our use of a serial recall of fixed length lists is functionally equivalent to a span measure. Finally, as predicted, the phonological coding measures correlated with serial recall performance.

However, there were some instances where outcomes were not in accordance with expectations. The search measures used in the present study were not correlated with recall although equivalent measures in the developmental literature do correlate with span. Inspection of the other relations involving the search measures suggests that they have not behaved as expected, even to the point of barely correlating with each other ($r = .22$). Given that they were chosen as markers of the search factor, a much stronger within-factor correlation was expected. It is very likely that the 20-second interval allowed for both tasks did not permit reliable measurement of search processes. In the absence of any reliability data, findings relating to these variables should be treated with caution.

Testing hypotheses about relations among variables

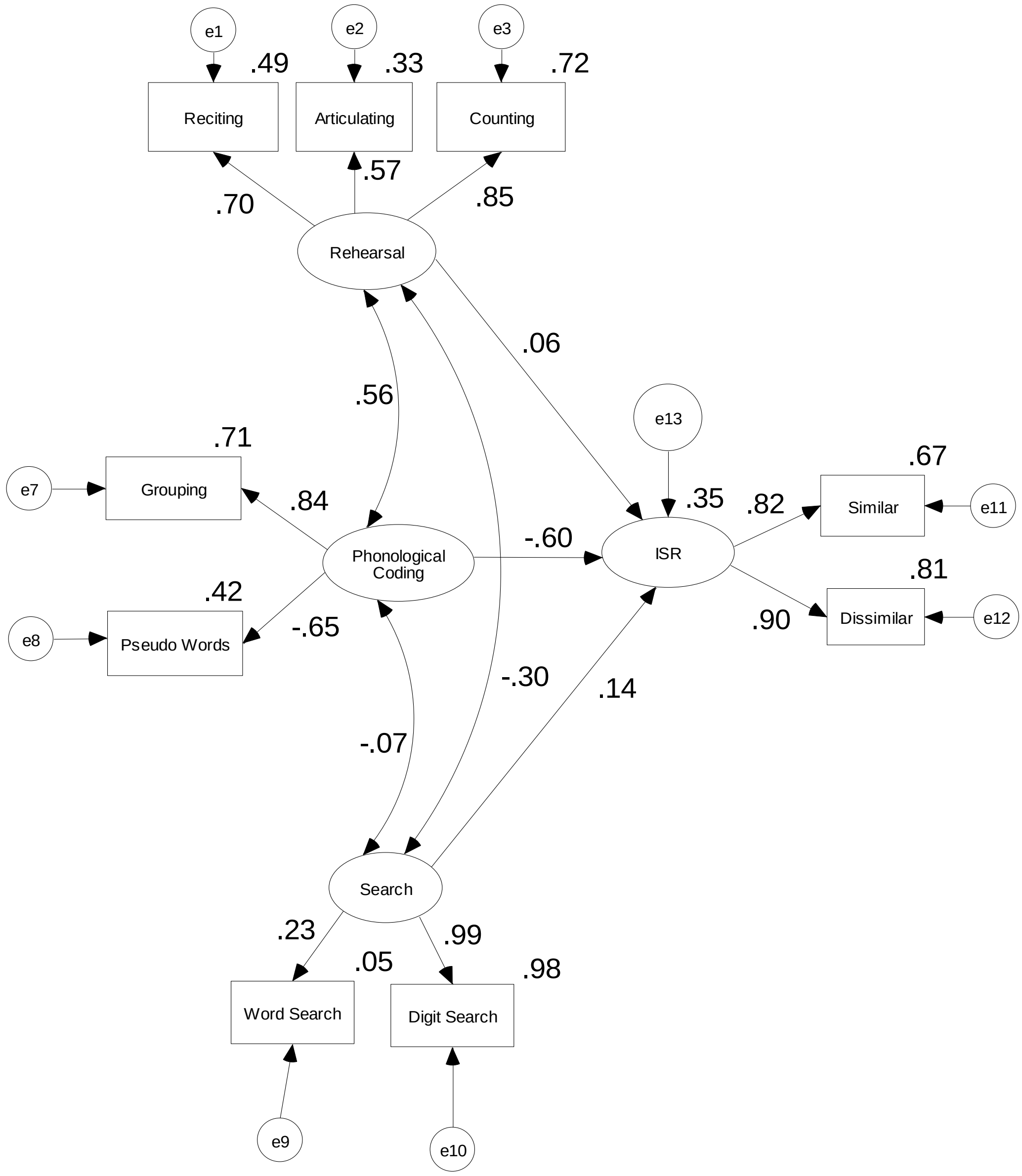
One simple test of the hypothesis that span depends more on lexical access than on rehearsal, phonological coding, or search processes is to check for significant differences among the correlation coefficients involving serial recall and the various groups of measures. The correlation coefficients in Table 1 are not independent so the normal Fisher’s Z transformation could not be used. Instead, Steiger’s (1980) Multicorr program was used. Multicorr performs a variety of complex correlational hypothesis tests, including tests of differences between within-sample correlation coefficients. Results indicated that the correlations involving Sound Grouping, the variable with the highest correlation with serial recall, were not significantly different at the .05 level from those between serial recall and lexical access or rehearsal measures. At this level, therefore, there was no support for the suggestion that lexical access has a stronger relationship with serial recall than other types of measure used here. In the type of design that we employed, however, the real interest is in the relations that exist among actual constructs, rather than among specific measures of these constructs. To investigate relations at the construct level, we employed structural equation modelling. These analyses are presented next.

The AMOS 4.01 statistical package (Arbuckle, 1999) was used on the covariance matrix generated by the twelve variables to test a full structural equation model. To select fit statistics appropriate for a maximum likelihood solution generated on a relatively small sample size, Hu and Bentler’s (1999) 2-index presentation strategy was followed. Using this strategy, the standardised root mean square residual (SRMR) and the comparative fit index (CFI) were chosen as the two indexes. Cutoffs of .08 for the SRMR and .95 for the CFI were taken to indicate good fit of the model to the data.

Following the design of the study, the first model described articulation time, counting time, and alphabet recitation as markers for a Rehearsal factor; lexical decision, non-word naming, and word naming as markers for a Speed of Lexical Access factor; sound grouping and pseudo words as markers for a Phonological

Coding factor; and word search along with digit search defining a Search factor. Recall of phonologically similar and dissimilar words formed two indicators of the Immediate Serial Recall (ISR) factor. In the structural part of the model, the first four factors were treated as correlated predictors of the ISR factor. Whilst the model explained 38% of the variance in ISR and fit statistics were very good (SRMR = .06; CLI = .99), a large correlation (r = .91) between the Lexical Access and Phonological Coding factors suggested that these two factors could either be combined or one of them could be dropped. A second model combining the factors was fitted to the data. The combined-factor model explained 32% of the variance in ISR and returned identical fit statistics. Two further models were tested, the first with Phonological Coding removed, the second with Phonological Coding present and Lexical Access removed. As expected, fit statistics were very good for both models (SRMR = .05; CLI = .99) but if the Lexical Access factor is omitted the explained variance in ISR decreases only to 35%. If the Phonological Coding factor is omitted instead of Lexical Access the explained variance drops much further to 26%. As such, although highly correlated, the two constructs do not appear to be equivalent. The model containing Phonological Coding but omitting Lexical Access is shown in Figure 1.

Figure 1. Measurement and structural model for predicting Immediate Serial Recall



Pathways from latent constructs to manifest variables were all significant (*p* < .05) except for the pathway from Search to Word Search and the pathways from Search and Rehearsal to ISR.

Setting aside for the moment the question of the relationship between phonological coding and lexical access processes, and with the proviso that the measures of search processes used in this study may not have been sufficiently reliable, we tentatively conclude that phonological coding processes are more important for explaining immediate serial recall than rehearsal or search processes. This is not to say that these other two processes are not involved in immediate serial recall. The technique employed in SEM is similar to standard regression wherein the beta weights can be interpreted as reflecting the <u>unique</u> contribution of the different constructs to the construct of immediate serial recall. To determine whether other variables are related, we need to go back to the table of correlations. Table 1 shows that all measures, except for Alphabet Recitation and the two Search markers, correlated with both serial recall measures. On the basis of these data, we conclude that whilst other variables may be related to serial recall performance, Lexical Access and Phonological Coding are the primary contributors. Further evidence is provided by the fact that if all other factors are removed from the model shown in Figure 1, leaving Lexical Access and Phonological Coding as the sole predictors, R-Square decreases by just 4%. If Lexical Access and Phonological Coding are removed and Rehearsal and Search left in the model, R-Square decreases to 11%. The most parsimonious model of all is one which includes just the Phonological Coding factor which, on its own, explains 34% of the variance in ISR.

Discussion

The current study built upon an earlier study by Tehan and Lalor (2000) and replicated their findings by showing that speed of lexical access accounted for much more of the variance in immediate serial recall measures than did articulation speed. The current results show that speed of lexical access extends beyond simple span measures of capacity to fixed length trials where open sets of words are used. The influence of speed of lexical access seems to be a strong determinant of performance on a variety of immediate serial recall tasks.

In addition to rehearsal and lexical access variables, the current study examined the effects of two other variables, facility with phonological coding and memory search speed. Neither of these variables behaved totally as expected. The degree of overlap between the Phonological Coding and Lexical Access factors was not anticipated and could indicate that these constructs may well be identical. Alternatively, the constructs may not be identical but overlapping in such a way that one subsumes the other. Phonological Coding is the more likely candidate for the over-arching role. For example, in the pseudohomophone naming task, participants may begin by consulting the lexicon to discover whether a match exists between the collection of letters that constitute the item and words in the lexicon before using phonological knowledge to sound out the word. This first part of the phonological coding task provides the link with lexical access, whilst the second part constitutes a unique coding process that creates significant additional overlap with immediate serial recall tasks. This additional variance explained by the phonological coding tasks does not challenge the importance of lexical access but suggests that there are other constructs that can take us further along the path to explaining variance in ISR tasks.

Our findings in relation to the Search construct are at odds with claims about the importance of search processes in the literature (Cowan et al., 1998). However, before making substantive interpretations of these different outcomes, we firstly wish to discuss the possibility that the different outcomes can be attributed to differences in the measures used across the different studies.

While there is surface similarity between the tasks used in the current study and those used by Cowan et al., (both search for letters in a large matrix) the search measures are quite different. In the current study we measured the time taken to search for a single target (or two targets) and used this simple time measure as the basis for our estimate of the search construct. Cowan et al. used a more complex measure of search speed, the slope of a regression line. It is thus clear that we have very different measures of the construct and the lack of a correlation with our simple measure may be attributable to the different way in which we have operationalised the search construct.

The Cowan et al. and the Hulme et al. approach of using regression slopes as a measure of search speed has face validity. However, it is important to reiterate that these measures are derived in situations where there is high memory load. Consequently, the correlation between slope and memory span may reflect the common memory component. We do not see any simple solution as to how one might calculate a search measure that has high face validity yet makes little demand upon memory. Our first approximation was to use the same sort of search task as those used in previous studies, but to minimise memory load.

A second possible additional reason for the discrepant findings is that our measures of search were not sufficiently reliable to allow robust correlations with other variables to emerge. The fact that measures of letter cancellation of the type used in the current study have been shown to have small but reliable correlations with memory span in established test batteries such as the Woodcock-Johnson Revised Tests of Cognitive Ability suggests that our own data may have underestimated the strength of this relationship. Following this same line of reasoning, however, we note that although the literature suggests a relationship between cancellation tasks of the type that we used and memory span measures, those correlations are not high and account for approximately only 10% of the variance in span measures. For larger relationships to emerge, it is probably necessary to induce a memory load component in the search task itself.

Implications for Current Models of Short-term Memory

In the working memory model, rehearsal is seen to be a prime determinant of individual differences in span performance. Tehan and Lalor (2000) found that measures of speeded articulation made only a modest contribution to span performance, compared to lexical access variables. The results of the current study are much the same in that there is a substantial contribution of lexical access and phonological coding variables but very little contribution from rehearsal measures. The data do not support the conclusion that rehearsal is a primary determinant of span.

The working memory model, and other models of immediate recall, assume that items in memory are represented in terms of their phonological characteristics. Our results confirm this basic assumption and suggest that individual differences in the ability to use phonological codes is an important contributor to individual differences in immediate recall. These results reflect similar findings in the developmental literature (Bowey et al., 1992; Wagener et al., 1997)

While many of the formal models of immediate serial recall propose some sort of redintegration procedure, none are developed to the extent that the structure and functioning of lexical memory are incorporated. Word versus non-word differences are readily explained by most models, but word frequency, word class and even sub-lexical effects (Nimmo & Roodenrys, 2002) are not readily explained. Given the importance of lexical access to immediate recall, this aspect of the various models will need to be developed if a complete account is to be realised.

Cowan et al's (1998) ideas tend to reflect traditional models of short-term memory that propose that items are resident in a limited capacity short-term store that can be directly accessed via a search process (Sternberg, 1966). That is, that retrieval involves a search through the contents of a limited set of items during pauses between the recall of items (Cowan et al., 1992). Thus, a clear prediction of the Cowan approach is that individual differences in speed of search through a limited set of items should be related to span. Our attempts to measure this process do not support Cowan’s assumptions, however, as we have suggested earlier, our measure may not be an adequate operationalisation of the search process.

In Cowan’s model speed of search is only one component. He argues that items rapidly lose their fidelity during the recall process. During pauses in recall, fading traces are updated and reactivated so that search can continue. Assuming that reactivation is much the same as what others label redintegration, our data suggests that individual differences in this part of the retrieval process are important, and may be more important than the speed of search itself.

Theoretical Caveats

The current research was aimed at testing ideas about how immediate recall works. These ideas were the basis for the directionality (and implied IVs and DVs) among constructs in our theoretical model. Moreover, in light of these ideas we interpreted the current results as being consistent with a redintegration process. That is, the high correlation between our phonological and lexical access variables combined with their predictive power is seen as supporting evidence for such a process.

However, as is the case with correlational designs involving latent variables, objections can be raised regarding firstly, the identity of the latent variables and, secondly, the direction of causality. We have argued that lexical decision, word naming and non-word naming are measures of a latent speed of lexical access construct. These are the tasks that have been used ubiquitously to explore repetition, phonological and semantic priming effects in lexical memory and as such we are confident that we have the best measures available for measuring speed of access to lexical memory. However, as one reviewer pointed out, these tasks are primarily visual word decoding tasks and what we may well be measuring is participants' facility in decoding visually presented letter strings. As such, the latent construct might be considered as speed of item identification at encoding. Tehan and Lalor (2000) considered this possibility and noted item identification could not be ruled out in their data, as it cannot be ruled out in the current data. However, they noted that lexical access speed correlated with pause times during recall. They argued that participants were accessing lexical memory during those pauses, presumably trying to redintegrate a degraded phonological representation.

The possibility that the latent construct reflects decoding ability has implications for the directionality of the relationships between constructs. There is a wealth of literature in the developmental domain indicating that those children who have difficulty on word decoding tasks also have problems with memory span and with phonological processing tasks. Some of this literature has explored the possibility that deficits in memory span may play a causal role in decoding ability. For instance, it has been proposed that beginning readers have difficulty in holding phonemes in memory while they are blended into words (Wagner et al., 1994) or alternatively, phonological representations must be held in memory for long-term learning of letter-sound correspondences (Gathercole, 1995). However, at this stage there is just as much contradictory evidence as supporting evidence and the latter deals primarily with beginning levels of reading. Thus, there is no definitive evidence indicating a causal role of span on decoding skill and on the basis of the developmental literature, it seems unlikely that differences in span would be the cause of the effects we have seen with adult participants.

In the developmental literature, the stronger evidence is for the proposition that problems in phonological coding may have a causal effect upon both word decoding and span because phonological representations underpin performance on both tasks. Again, the current research cannot discriminate between these options. The research does, however, confirm that these relationships are present in an adult population (Bowey et al., 1992; Wagner et al., 1994, 1997)

Implications for individual differences psychology

Although short term memory tasks have been part of the intelligence tradition since the first Binet/Simon scales, they have not attracted a great deal of attention in the individual differences literature. Carroll's (1993) seminal work on human cognitive abilities, for example, mentioned only a handful of studies on memory *per se* with most research including memory among a batch of other ability constructs. Whilst studies of the place of memory within the structure of abilities are important for deciding among theoretical models of intelligence, they do not highlight particular processes that underlie particular abilities. Guided by theoretical developments in experimental cognitive psychology, research of the type demonstrated in the present studies is therefore important for identifying processes that underpin broader cognitive constructs, in this case short term memory. The demonstration that lexical access and phonological coding processes consistently make the strongest contributions to short term memory performance should alert individual differences researchers to the need to delve beneath some of the broad constructs that are typically the focus of research in this tradition. As Ackerman and colleagues noted, previous research emphasising the role of constructs such as general intelligence and reasoning has tended to overlook the role of contributing processes, such as speed of processing (Ackerman, Beier, & Boyle, 2002). The fact that the current set of theoretically-derived measures was able to account for between 32 and 38 percent of performance in span indicates the benefits that can be gained from cross-fertilisation between the experimental and individual differences fields.

Conclusion

Our results confirm current thinking that redintegration by way of lexical access from phonological codes is essential processes involved in immediate memory. Individual differences in such abilities produce corresponding individual differences in immediate serial recall. We have found much less of a role for other supposedly crucial processes like rehearsal speed or speed of search through short-term memory.

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.

Arbuckle, J. L. (1999). *Amos for Windows. Analysis of moment structures. Version 4.01.* Chicago, IL: Smallwaters Corporation.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18A,* 362-365.

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

Baddeley, A.D, Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour, 14,* 575-589.

Bowey, J.A., Cain, M.T., & Ryan, S.M. (1992). A reading-level design of phonological skills underlying fourth-grade children's word reading difficulties. *Child Development, 63*, 999-1011.

Brown, G. D. A., & Hulme, C. (1995). Modelling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language, 34*, 594-621.

Burgess, N. & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551-581.

Caplan, D., Rochon, E., & Waters, G. (1992). Articulatory and phonological determinants of word length effects in span tasks. *Quarterly Journal of Experimental Psychology, 45A*, 177-192.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Chen, C., & Stevenson, H. W. (1988). Cross-linguistic differences in digit span of preschool children. *Journal of Experimental Child Psychology,46*, 150-158.

Coltheart, M. (1980). Reading, phonological recoding and deep dyslexia. In M. Coltheart, K. Patterson & J. C. Marshall (Eds.) *Deep Dyslexia* London: Routledge & Kegan Paul.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake and P. Shah (Eds.) *Models of Working Memory: Mechanisms of active maintenance and executive control.* (pp. 102-134). New York: Cambridge University Press.

Cowan, N., Day, L., Saults, J.S., Keller, T.A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language, 31,* 1-17.

Cowan, N., Keller, T.A., Hulme, C., Roodenrys, S., McDougall, S. & Rack, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory and Language, 33,* 234-250.

Cowan, N., Wood, N. L., Wood, P. K, Keller, T.A., Nugent, L. D & Keller, C. V. (1998). Two separate verbal processing rates contributing to short-term memory span . *Journal of Experimental Psychology: General, 127,* 141-160.

Ekstrom, R. B., French, J. W., & Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Farrell, S. & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9,* 59-79.

Ferguson, A.N., Bowey, J.A., & Tilley, A. (2002). The association between auditory memory span and speech rate in children from kindergarten to sixth grade. *Journal of Experimental Child Psychology, 81*, 141-156.

Friendly, M., Franklin, P., Hoffman, D., & Rubin, D. (1982). The Toronto Word Pool: Norms for imaging, concreteness, orthographic variables, and grammatical usage for 1080 words. *Behavior Research Methods and Instruments, 14,* 357-399.

Gathercole, S.E. (1995). Is nonword repetition a test of phonological memory or long term knowledge? It all depends on the nonwords. Memory and Cognition, 23, 83-94.

Hatcher, P. J. & Hulme, C. (1999). Phonemes, rhymes, and intelligence as predictors of children's responsiveness to remedial reading instruction: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 72,* 130-153 .

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling, 6*(1), 1-55.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology* 38, 241-253.

Hulme, C, Maughan, S., & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C., & Roodenrys, S. (1995) Practitioner review: Verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry, 36*, 373-398.

Kail R. (1997). Phonological skill and articulation time independently contribute to the development of memory span. *Journal of Experimental Child Psychology, 67*, 57–68.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory and Cognition, 18*, 251-269.

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

Nimmo, L. M. & Roodenrys, S. (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics, 23,* 643-659.

Page, M. P. A. & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761-781.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory and Cognition, 21*, 168-75.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning Memory and Cognition, 12*, 419-425.

Snowling, M. & Hulme, C. (1989). A longitudinal case study of developmental phonological dyslexia. *Cognitive Neuropsychology, 6*, 379-401.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245-251.

Sternberg, S. (1966). High speed scanning in human memory. *Science, 153*, 652-654.

Tehan, G., & Humphreys, M. S. (1988). Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note. *Bulletin of the Psychononmic Society, 26*, 293-296.

Tehan, G., & Humphreys, M.S. (1995). Transient phonemic codes and immunity to proactive interference. *Memory and Cognition, 23,* 181-191.

Tehan, G., & Lalor, D.M. (2000). Individual differences in memory span: The contribution of rehearsal, access to lexical memory, and output speed. *The Quarterly Journal of Experimental Psychology, 53A*, 1012-1038.

Thurstone, L. L. (1936). The factorial isolation of primary mental abilities. *Psychometrika, 1*, 175-182.

Turse, P. L. (1940). *Turse shorthand aptitude test.* Yonkers-on-Hudson: World Book Co.

Wagner, R. K., Torgesen, J. K. &. Rashotte C. A. (1994) Development of reading-related phonological processing abilities: Evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73–87.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., Hecht, S.A., Barker, T.A., Burgess, S.R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading a children develop from beginning to skilled readers: A five-year longitudinal study. *Developmental Psychology, 33*, 468-479.

Walling, J. R., McEvoy, C. L., Oth, I E., & Nelson, D. L. (1994). The University of South Florida rhyme category-norms. Unpublished manuscript, University of South Florida.