Individual Differences in Memory Span: The Contribution of Rehearsal, Access to Lexical Memory and Output Speed.

Gerald Tehan

David M. Lalor

University of Southern Queensland

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

This is the authors' final corrected pre-publication version of this paper.
Accessed from USQ ePrints http://eprints.usq.edu.au/archive/1978/

Mailing Address:

Gerry Tehan
Department of Psychology
University of Southern Queensland
Toowoomba, 4350
AUSTRALIA
Phone number: + 61 7 4631 2376

e-mail: tehan@usq.edu.au

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 subject 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 explore individual differences in span with respect to measures of rehearsal, output time, and 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 span, than either rehearsal speed or output factors.

Individual Differences in Memory Span: The Contribution of Rehearsal, Access to Lexical Memory and Output Speed.

The memory span task is perhaps the most widely used measure of short-term memory performance. At face value, recalling a short string of items in order appears to be a very simple task. However, recent research suggests that the task is far from simple. Span is influenced by numerous speech-based variables (Baddeley, 1986), by variables involved with the maintenance of serial order (Estes, 1972; Henson, Norris, Page & Baddeley, 1996), and by variables that reflect lexical access (Hulme, Maughan & Brown, 1991; Poirier & Saint-Aubin, 1995). Furthermore, these speech-based, lexical and order variables all interact with each other. Such complexity has ensured that the memory span task continues to play a pivotal role in the development of theories of short-term recall.

In the current experiment we are interested in examining individual differences in three areas that are argued to be strong determinants of span: rehearsal speed (Baddeley, 1986; Baddeley, Thomson & Buchanan, 1975), forgetting during list recall (Cowan, 1992; 1993) and the ability to use long-term lexical knowledge to reconstruct degraded short-term traces (Brown & Hulme, 1995; Hulme, Maughan & Brown, 1991; Schweickert, 1993). In the span research that we know of, no one has addressed these issues in the same experiment. To preview the results somewhat: we replicate the finding that rehearsal rate and span are strong correlates of each other; we show that output differences contribute to individual differences in span; and finally, we show that prototypic lexical memory tasks share common variance with span performance.

Memory Span and Rehearsal Rate

Rehearsal has long been postulated to play a critical role in span (Atkinson & Shiffrin, 1968). However, we want to focus upon its role in what has come to be known as trace decay plus rehearsal models (Baddeley, 1986; Schweickert & Boruff, 1986; Schweickert, Guentert & Hersberger, 1990), and more specifically in Baddeley's (1986) concept of the articulatory loop. It is now well established that span is affected by such speech-based variables as phonemic similarity (Conrad & Hull, 1964; Baddeley, 1966), word length (Baddeley et al. 1975), articulatory suppression (Peterson & Johnson, 1971), and unattended speech (Salamé & Baddeley, 1982). Baddeley has proposed a two-component structure to account for the interactive effects of these variables on span. The first component is a rapidly decaying phonological store in which phonological representations of to-be-remembered items are temporarily held. The second component is a verbal rehearsal process that has the dual function of translating visual input into phonemic representations and maintaining phonemic representations in a readily retrievable state. The rehearsal process is seen to be responsible for producing the word length and suppression effects as well as being the limiting factor on span performance. That is, the bottleneck for span performance is related to the time needed to rehearse items (Baddeley et al., 1975; Cowan et al., 1992; Hulme, Thomson, Muir, & Lawrence, 1984; Naveh-Benjamin & Ayres, 1986; Schweickert & Boruff, 1986; Standing, Bond, Smith & Isely, 1980; Stigler, Lee, & Stevenson, 1986).

Empirical support for such an assertion has its origins in Baddeley et al.'s (1975) demonstration that memory span was related to the spoken duration of the material. Words that took longer to pronounce were associated with a lower level of recall. In addition, Baddeley et al. indicated that a positive linear relationship existed

between the number of items that could be recalled and the rate of rehearsal for the particular material, where rehearsal was measured using speeded reading and articulation tasks. The former task involved the reading of a list of 50 words as quickly as possible, while the latter rehearsal estimate involved the continuous repetition of a list of three words. The regression equation involving span and rehearsal rate suggested that span could be predicted on the basis of the number of items that a subject could either read or articulate in approximately 2 seconds.

Span-rehearsal rate correspondences have subsequently been used to account for many of the structural, developmental, cross-cultural and individual differences that are observed in span performance. Thus, Schweickert and Boruff (1986) demonstrated that span across a variety of materials (e.g., digits, words, colours, shapes, consonants, and nonsense syllables) was matched by differences in reading rate for the same materials. Developmental differences in span are accompanied by corresponding differences in articulation rate (Hulme et al., 1984). Cross-cultural differences in digit span also appear to be predictable on the basis of how quickly the digits are pronounced in the various languages (Chen & Stevenson, 1988; Elliott, 1992; Ellis & Hennelly, 1980; Naveh-Benjamin & Ayres, 1986; Stigler et al., 1986). For present purposes the most important aspect of span-rehearsal rate correspondence is that individual differences in adults' memory span can be predicted by differences in rehearsal rate (Baddeley et al., 1975). That is, those subjects who rehearse fastest also tend to have the highest digit span.

In summary, the articulatory loop and similar trace decay models suggest that memory span is determined by rehearsal speed. Given a memory trace that decays in about two seconds (Schweickert & Boruff, 1986; Schweickert et al., 1990), span is equivalent to the amount of material that can be kept active via rehearsal for this length of time. It follows that those subjects who can rehearse the fastest can keep more items in an active state and thus produce higher spans.

Alternatives to Rehearsal: Item Identification

The notion that rehearsal is the primary determinant of span performance has not gone unchallenged. An alternative view regarding the relationship between articulation rate and memory span involves the mediation of a third variable. In an extensive review of the developmental literature, Dempster (1981) examined 10 potential sources of individual and developmental differences in span, including, rehearsal, grouping, chunking, retrieval strategies, item identification, item ordering, capacity, interference susceptibility, search rate, and the output buffer. He concluded that the variable most likely to contribute to individual and developmental differences in memory span was item identification time. Furthermore, he suggested that speed at identifying the names of words was an integral part of reading, thus suggesting that span-reading rate correspondences were underpinned by the same process.

Case, Kurland, and Goldberg (1982) tested the item identification speed hypothesis with respect to the developmental increase in span. In their first experiment, they found a linear relationship between increases in identification speed (where the time taken to name a visually presented word was used as a measure of identification speed) and increases in memory span (r =.74), for subjects between the ages of three and six years. In a second experiment, groups of adults and six-year-old children were equated on identification time for nonwords and words, respectively. The identification time measure for this experiment again involved the time it took to begin the naming response to visually presented words. Case et al. showed that by lowering the familiarity of the stimulus items for adults, measures of naming latency

and span for the respective items did not differ significantly between the groups. They therefore suggested that a causal relationship exists between these variables (however, see Henry and Millar, 1991 for a different conclusion).

Item Identification Versus Rehearsal Speed

Hitch, Halliday and Littler (1989; 1993) tested the rehearsal-based articulatory loop model and the identification time hypothesis with respect to span across different age levels. The 1989 study involved a comparison between groups of 8- and 11-year-old children on span for one-, two- and three-syllable words; span under articulatory suppression; rehearsal rate; and item identification time. Measures of rehearsal rate included articulation rate (i.e., continuous rapid repetition of word-pairs) and reading rate in which subjects had to read lists of 16 words aloud as quickly as possible. Two estimates of identification time were employed, the time to name visually- and auditorially-presented words. The results revealed that span under rehearsal conditions was a linear function of rehearsal rate (as reflected by either reading or articulation rate) across differences in both word length and age. Contrary to the predictions of the identification time hypothesis, there was no evidence that the word length effect was mediated by differences in identification time. That is, although longer words were subject to poorer recall, no significant difference was found between identification times for long and short words. Under suppression the word length effect was eliminated, as per expectation, but suppression did not eliminate the age differences in span, as would be expected if rehearsal differences were the sole determinant of span. Furthermore, span under suppression conditions was highly correlated with identification time, but was not significantly related to the rehearsal rate measures, both across and within age groups. Hitch et al. interpreted these results as being strongly suggestive of the independent contribution of item identification processes.

At an empirical level, the item identification literature has indicated that naming latency is a correlate of span at least for young subjects under some conditions. We are interested in determining if naming latency is a correlate of span in an adult sample. At a theoretical level the results of these studies suggest that item identification is distinct from rehearsal speed. However, both processes are seen to have their effect during learning a list. Rehearsal is seen to maintain items that have been encoded and item identification is seen to reflect the facility with which items can be encoded in the first place (Dempster, 1981). We have some reservations with this notion that naming latency reflects ease of encoding. We hope to demonstrate later that naming latency may well reflect facility with retrieval, a feature of some of the more recent models of span (Brown & Hulme, 1995; Schweickert, 1993).

Rehearsal Plus Reconstruction

Despite the initial appeal of the rehearsal speed explanation, recent evidence suggests that it does not offer a complete account of memory span performance. Instead, a number of studies indicate that access to lexical memory at retrieval contributes to span performance (Hulme et al., 1991; Poirier & Saint-Aubin, 1995; Schweickert, 1993). Lexical features of verbal materials often have a similar impact upon span under rehearsal and articulatory suppression conditions. Several studies have suggested that word-frequency contributes to span performance, and that this contribution is not mediated by articulatory loop processes (Roodenrys, Hulme, Alban, Ellis, & Brown, 1994; Tehan & Humphreys, 1988). Word class is another such variable with content words (e.g., nouns, verbs, and adjectives) being recalled better than function words (e.g., prepositions, pronouns, articles) under both rehearsal and

suppression conditions (Bourassa & Besner, 1994; Tehan & Humphreys, 1988).

Results such as these have led a number of researchers to theorise that 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). For example, Hulme et al. (1991) demonstrated that span for nonwords is lower than that for words, although both show standard word length effects. Furthermore, when span for different lengths was plotted as a function of the time it took to pronounce the items, linear relationships were obtained for both words and nonwords and the slopes were equivalent. The nonword function, however, exhibited a lower intercept. The corresponding differences in articulation rates were not observed. That is, words and non-words were articulated at the same rate, a finding which is inconsistent with the position that span is based solely upon rehearsal rates. Hulme et al. suggested that their results support a two-component model of individual differences in span performance. The equivalent slopes for the span-articulation function was interpreted as a contribution of the articulatory loop (via differences in rehearsal rate). They assumed that the lower intercept for the non-word function reflected the absence of LTM lexical representations of the non-words. In their second experiment, Hulme et al. provided causal evidence in support of a LTM contribution. In this experiment memory span for groups of Italian words was determined both before and after the subjects learned the English translations of the words. As predicted, this manipulation of familiarity via the creation of LTM representations resulted in a significant increase in span in the post-learning phase. Again, this increase could not be accounted for in terms of an increased speech rate.

With respect to the nature of the information in LTM that contributes to span performance, Hulme et al. argued that it is the phonological representations of the material that are crucial. That is, although the operation of the articulatory loop is insensitive to the lexical status of the stimulus material, long-term phonological information could be accessed to facilitate the reconstruction of degraded phonological representations. Hulme et al. indicated that the magnitude of the LTM contribution to span might be predicted from factors such as familiarity and ease of retrieval of information from LTM. This line of reasoning suggests that pattern (word) completion ability following the retrieval of incomplete traces may be a possible source of individual differences in span performance.

In summary, these studies suggest that LTM, in the form of lexical information, is used to supplement the representations of stimulus items in the phonological short-term store. The Hulme et al., account provides two avenues to explore individual differences. Firstly, there is the potential for individual differences in the speed of lexical memory access to contribute to individual differences in memory span. Secondly, the Hulme et al. (1991) redintegration process is essentially a pattern completion process. There could well be a relationship between span and pattern completion ability.

Recall/Output Processes

The articulatory loop explanation primarily involves processes that are occurring during the study of a list. That is, there is a trade-off between rehearsal and decay prior to recall. The redintegration approach suggests that speed of access to lexical memory at retrieval might well be an important component of span performance. Cowan (1992; 1993) has suggested another possible influence: the rate of forgetting during recall. He has argued that span is determined largely by the

efficiency of retrieval and reactivation of material during the response period. Empirical support for the significance of recall characteristics had been presented earlier by Cowan et al. (1992). These authors indicated that word length effects were due to output delays. They found that this effect was most pronounced for earlier serial positions. That is, the recall of longer words at the beginning of a list caused more decay of the remaining items in the list than if the words at the beginning were relatively short.

Cowan (1992) subsequently investigated the speech timing characteristics of the memory span responses of 4-year-olds. Separate timing measures were taken of pronunciation of individual words, the preparatory interval (i.e., from the end of the stimulus list to the beginning of recall), and the interword pauses in the response. Cowan found a positive association between span and the duration of spoken recall. Furthermore, more capable subjects responded at the same rate as subjects with lower spans, but the responses continued for a longer period. A positive relationship between the interword pause times and list length was also found. Separate analyses involving span-length versus subspan-length responses revealed that the interword pauses for subspan-length lists were significantly shorter.

Cowan et al. (1994) also produced results that are relevant to the interpretation of the output measures. Specifically they found that the spoken duration of words during output was affected by word length, but not by age. However, the pause times were affected by age, but not word length. To explain the output data, Cowan argues that pause times reflect retrieval and reactivation efficiency. That is, in the pauses between items subjects are assumed to retrieve and refresh the unrecalled items, with high-span and/or older children being more proficient at this than low-span/younger children. Word durations on the other hand, reflect periods during which non-recalled items decay.

In summary, spoken recall studies suggest that list items are subject to rapid decay during recall. Subjects engage in a covert retrieval process during interitem response pauses in order to retrieve the next item and to reactivate the memory traces of the remaining list items. Individual differences in span are related to differences in the efficiency with which this process is undertaken. The question that we are interested in is that while this may be an adequate explanation for children's span, it may well not apply to adults, in that retrieval may well be automated. There is evidence that adults may not show this pattern, Stigler et al., (1986) looked at output timing in adults and found that span response averaged at about 2 seconds. This result has often been interpreted as output time being fixed for adults.

Experiment 1

The above review has attempted to make the point that rehearsal, item identification, long-term memory access, and verbal output variables all make a significant contribution to span performance, at least with children as subjects. In the current experiment we first want to establish that these variables are characteristic of adult span performance. We then hope to show that by using an individual differences methodology, we can establish the relative contribution of these variables to span. However, before proceeding to the experiment we wish to give a rationale for the measures that we have utilised.

Rehearsal Measures.

The traditional measures that have been used to argue for a primary role of rehearsal are a reading task, where subjects are asked to read a list of words as quickly as possible; and an articulation rate task, in which subjects are asked to repeat

the same two or three words as quickly as possible for a set number of times. Although we have adopted these tasks, we have some reservations about what these tasks actually measure. For instance, the reading task looks very much like a continuous word naming task. Given the evidence that item identification is a correlate of span, the reading task may very well measure speed of item identification rather than rehearsal rate. Moreover, Naveh-Benjamin and Ayres (1986) have indicated that speed of reading is not necessarily the key facet of the reading rate - span correlation. They found that the correlation was just as strong with normal reading rate as with speeded reading.

The articulation measure, while probably having more face validity as a rehearsal measure, is not without problems either. Chase, Lyon and Ericsson (1981) have suggested that the correlation of span with articulation rate is an artefact of the fact that memory for order is involved in both tasks. Their reason for this assertion is that the correlation only holds when the number of words that have to be articulated approaches span. They report 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.

In contrast to the Chase et al. (1981) findings, Hulme et al. (1984) measured 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. Thus, for children at least, articulation rate of a single item was a reliable predictor of span. Similarly, Hitch et al. (1989) found that rapid articulation of word pairs correlated with span. The Chase et al. findings may be questionable.

Given that some doubt hangs over the two most popularly used measures of rehearsal, we supplemented these with two other tasks that measure articulation speed but should not have any large memory component associated with them. That is, we measured how long it took subjects to articulate two overlearned sequences: the letters of the alphabet and the numbers from 20 to 40. Neither of these sequences should engender high memory loads and access to these items in long-term memory should be highly automatic.

Output Measures

We have utilised identical measures to those used by Cowan (1992). These are described in more detail in the Methods section.

Long-term Memory Measures

Part-word Completion. Those models that argue for a long-term component to span generally assume that a word must be identified from a partially degraded phonemic trace (Brown & Hulme, 1995; Schweickert; 1993). We sought to try and capture this redintegration process by measuring the time it took people to generate a word given a part-word stimulus. In the first instance we presented word fragments (e.g. ar_ow) and asked them to complete the word. In the second instance we presented subjects with a word ending (e.g. _and) and asked them to generate three instances of words that had that particular ending. One problem with the latter measure is that word endings function as efficient retrieval cues in long-term memory and ending categories tend to have many of the same attributes as taxonomic categories (Nelson & McEvoy, 1979). In order to check on just what we were measuring with the ending completion task we also required subjects to produce three instances of taxonomic categories (e.g., Animals) as well. This also allowed the possibility that speed of access to categorical information might play a part in span.

Lexical Access. In the introduction to this paper naming latency was identified as a correlate of span, at least for children. We have included this test in our

battery to see if naming latency is a predictor of individual differences in span for an adult sample. We have also measured latency to pronounce non-words. 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). In the dual route models, lexical access for visually presented words can take place either directly from orthography or indirectly from orthography, through phonology to lexical memory. With non-words the indirect route is the only one available. Given the obligatory nature of phonological information with non-words, non-word naming may well be a more sensitive test of the redintegration process than word naming.

The correlation between span and naming latency has prompted us to also measure performance on the lexical decision task. The naming task is one of the three standard tasks that is used to explore visual word recognition processes in lexical memory, the others being perceptual identification and lexical decision. If naming latency reflects lexical access processes, rather than perceptual fluency, other lexical access tasks should also correlate with span.

With regards to the expectations of the study, the only firm expectation was that we should be able to replicate the relationship between span and rehearsal measures. If factors that affect children's span are also determinants of adult span then we should see the effects of output processes on span. If lexical access is important a significant relationship between span and lexical decision, word naming and non-word naming should emerge. If pattern completion plays a role the part-word completion tasks might well be correlated with span.

Method

Subjects

A total of 50 undergraduate students (11 male, 39 female) from the University of Southern Queensland participated in this study. All subjects received course credit in return for their participation. Three subjects had trouble recalling six digit lists and consequently some of the analyses are based upon the performance of the remaining 47 subjects.

Measures and Procedure

Subjects were tested individually in a single session of approximately 30 minutes duration. All tasks were visually presented and all responses were audio-taped. Half of the subjects did the tasks in the following order: digit span, reading rate, articulation rate, alphabet recitation, counting, fragment completion, ending completion, category instance generation, lexical decision, word naming, and nonword naming. The remaining subjects completed the tasks in an identical sequence, with the single exception of digit span, which was completed last (in order to counterbalance the potential influences of anxiety and fatigue on this task).

Digit span. This task consisted of 16 lists of digits, which were presented visually, on a computer screen, at a rate of one digit per second. The 16 lists were divided into sets of four, with the first set comprising six digits. The lists for succeeding sets were increased by one digit, so that the final set consisted of four 9-digit lists. Each list was preceded by a "Ready" signal. The lists were constant across subjects and were initially selected randomly without replacement from the set of digits, 0-9. Subjects were instructed to repeat the digits aloud as they appeared on the screen, and to recall the items in correct order, immediately following list presentation. A "?" appeared on the screen directly following the final digit in each

list in order to signal to the subject that the recall period had commenced. Subjects were allowed unlimited time to verbally recall the digits. The responses for each list were tape recorded for later analysis of the output.

Digit span was calculated in the traditional manner (Hulme et al., 1991). That is, span was the longest list length at which all lists in a set, and previous sets, were recalled correctly, plus 0.25 for each subsequent correctly recalled list.

The timing characteristics of the verbal output of the longest list that participants could recall correctly were analysed using the SoundEdit waveform editor. Three measures were determined using this procedure: Preparation time, Pause time and Speaking time. Preparation time was the period from the end of the subject's verbalisation of the final digit during presentation to the beginning of the recall response. Pause time measured any periods of silence between digits in the output protocol. Speaking time refers to the amount of time that the subject was actually vocalising during recall of the sequence. Speaking time and pause time in combination would reflect the actual response time for the list.

Reading rate. The materials for this task included a list of 50 low-frequency words from the Toronto Word Pool norms (Friendly, Franklin, Hoffman, & Rubin, 1982). The low frequency criterion constituted a Kucera and Francis (1967) frequency count of less than 20. The words were typed in two columns on a single page and subjects were instructed to read the list aloud, as quickly and as clearly as possible. Reading was to commence as soon as possible following a "go" signal, given by the experimenter. Reading time was calculated in the same way that Speaking time was measured in the output of span lists.

Articulation rate. Subjects were required to repeat three words, 10 times in rapid succession, following the start signal. Articulation time represents the amount of time the subject was verbalising during the task.

Alphabet. Subjects were instructed to say the letters of the alphabet as quickly and as clearly as possible following the signal to begin. Again the time spent verbalising served as the dependent measure

Counting. This task involved speeded verbal counting from 20 to 40, inclusive, following the start signal. Counting time was measured in the same way as the other rehearsal measures.

Fragment Completion. The materials for this task included a list of 20 randomly selected 4- and 5-letter words from the Toronto Word Pool Norms. A single letter was then randomly deleted from the interior of each word. The resulting letter strings were presented individually on a computer screen in standard upper- and lower-case characters, with a "_" inserted in the space previously occupied by the deleted letter. Each subject's task was to verbally generate an English word from the letter string (e.g., ar_ow). Subjects were encouraged to respond as quickly as possible. We took both latency and correct solution measures, but because subjects found the task fairly difficult, the measure we have used is the number of solutions generated rather than the latency measure.

Ending Completion. Ten letter strings (comprising two or three letters) were displayed sequentially on the computer screen in upper-case characters (e.g., ENT). Subjects were required to generate three English words that ended in the particular string, as quickly as possible. We took latency measures for all three instances generated but the time taken to generate the third instance appeared to be the most sensitive and it is this measure that we use.

Category Instance. Ten category labels were selected randomly from the

norms provided by Battig and Montague (1969). Each category cue was presented sequentially on the computer screen, and the subjects were instructed to verbally generate three instances of the particular category as quickly as possible. Again, latency to produce the third instance appeared to be the most sensitive measure.

Lexical Decision. The materials for this task included 40 low-frequency 5-letter words from the Toronto Word Pool Norms. The letters in 20 of these words were subsequently rearranged in order to create nonwords, which were phonotactically legal. The newly created nonwords were then randomly interspersed among the remaining words. The list of words/nonwords was displayed sequentially on the computer screen at a rate of one item every 5 seconds. The subject's task was to verbally decide (i.e., "yes" or "no"), as rapidly as possible, whether the item was a legitimate English word. Lexical decision time was measured by measuring the time from the presentation of the word to the beginning of the response.

Word Naming. The stimulus items for this task included 20 low-frequency five-letter words from the Toronto Word Pool Norms. These items were displayed on the computer screen at a rate of one item every 5 seconds. The subjects were required to name these words aloud as quickly as possible upon presentation. Naming time was measured from the time the word was presented to the time when the subject first began to name the item.

Nonword Naming. The 20 nonwords for this task were constructed in a similar manner to that described previously with respect to the lexical decision task. The procedure was similar to that involved for the word-naming task. That is, the nonwords were displayed visually at a rate of one item every 5 seconds and subjects were instructed to verbally repeat the items as rapidly as possible.

Results and Discussion

Descriptive statistics. The means and standard deviations of all measures together with their correlations with both measures of span are presented in Table 1. In all situations where statistical significance is being asserted, an alpha level of .05 has been adopted.

Preliminary Analyses: Bench Mark Results.

Rehearsal Rates. As can be seen in Table 1, significant correlation coefficients were found between span and the traditional rehearsal rate measures. We replicate the Baddeley et al. (1975) results in showing a correlation of -.50 between span and reading rate and a correlation of -.47 between span and articulation rate. The values of these correlations are not all that dissimilar to those obtained by Baddeley et al.

Item Identification. The correlation between span and word naming that has been observed in young children appears to generalise to an adult population, in that a correlation of -.40 between span and word naming time is present. Furthermore, non-word naming also appears to be a correlate of span in an adult population.

Response Output characteristics. In Cowan (1992) the response output characteristics of the children in that sample and their correlations with span were as follows. The longest list correctly recalled was significantly correlated to speaking time, but not to pause time or preparation time. As is evident in Table 1, the same relationships are observed in the adult sample: span does not correlate with pause time or preparation time of the longest list correctly recalled, but does correlate with speaking time. The implication of this latter correlation is that those with higher spans spoke for the longer periods. This finding verges on the tautological, however, it is still important because if trace decay occurs at a fixed rate as the trace decay models

suggest, then high span participants should speak faster than low span participants. That is, the spoken duration of each word in the list for the high span participants should be significantly less than that of the low span participants. We tested this notion by dividing our sample into three groups: 15 people for whom a six digit list was the longest they could recall correctly, a group of 15 who could remember a maximum of seven digits and a group of 17 who could recall eight or nine digit lists. Mean speaking time for these groups are presented in the top row of Table 2. The differences in speaking time were reliable, F (2,44) = 14.68, MSE = .24. The low span group spoke for less time than the medium span group, F (1,28) = 7.72, who in turn spoke for less time than the high span group, F (1,30) = 7.50. However, average speaking duration per word was .32, .34 and .36 seconds for the low, medium and high span groups respectively, F <1. In other words, everyone spoke at the same rate but those with high spans spoke for longer. These findings are inconsistent with assumptions that immediate recall is based upon a temporally limited store of two seconds duration.

Cowan (1992) also looked in detail at the response timing characteristics of the longest list that subjects could recall correctly (maximal span) versus subspan lists. In Cowan's sample, the average spoken duration of each word remained relatively constant across all conditions but pauses were shorter in subspan lists than in maximal span lists. In Table 2 speaking time, word duration, preparation time and pause times on the longest list that people could recall accurately (n) and on lists that are one digit shorter than the maximal span lists (n-1) are presented. As already indicated there were reliable differences in speaking time, but not in word duration for the low, medium and high span groups in the maximal span lists. For the medium and high span groups, there was also a reliable difference in speaking time between span and subspan lists, F (1,26) = 56.28. For preparation times there were no differences across span groups or between span and subspan lists.

Pause times were extremely variable as is evident by the standard deviation in Table 1. Approximately a third of the subjects did not produce any pause during the output of the longest list they could correctly remember and this proportion increased for subspan lists. At the other extreme there were four subjects who had pauses between two and five seconds before correctly the longest list. These latter subjects took long pauses before correctly recalling the last item in the list whereas most other subjects once they could not recall the final item immediately, gave up and were scored as incorrect. In the following analysis of pause times we have eliminated the scores of these four multivariate outliers in order to reduce the amount of variability in the pause times. Consequently, for the span lists there were no differences in pause times between the low, medium and high span groups, F (2,40) < 1, nor was there any difference between the medium and high span groups on the subspan lists, F (1,18) < 1. However pause times were significantly longer on the span lists than on the subspan lists, F (1,26) = 27.39. The data indicate that as subjects approached span length lists, longer pauses became evident and this was true for low, medium and high span groups.

In sum, the pattern of response characteristics of the adult sample were identical to the 4 year-old sample. Those with high spans did not differ from those with low spans on preparation time, pause time or the rate at which they pronounced individual words. The only difference was that the higher span groups recalled more items and this extended the time for which they were speaking. Finally, pause times were equivalent for all groups at maximal span. These pauses were both less frequent

and shorter for subspan lists than for span lists. We will return to the significance of these pause times a little latter.

Data Reduction and Regressions

The benchmark analyses indicated that the relationships between span and rehearsal, naming and output variables that have been observed in other research are also found in the present data set. The more exploratory analyses of our data first involved a factor analysis to identify the relationships that exist between the non-span measures. We then attempt to predict digit span on the basis of these latent variables.

The measures that we selected were aimed at providing complementary indices of particular constructs. We first looked at the correlations among variables to see if the data supported the way in which we conceptualised the measures. These correlations are presented in Table 3. In looking at the four rehearsal measures, articulation time, counting time, alphabet naming time and reading time, correlations are reasonably strong with an average correlation amongst the four measures of .47. The correlations among the four lexical access measures, word and non-word naming and word and non-word lexical decision times, are also strong, averaging .77. The set of measures that appears to share the least variance is the word completion variables of ending completion, category instance and fragment completion tasks. The average correlation amongst the three variables is only .32 and of more concern is the fact that the fragment completion and ending completion task, both of which were selected because they involve the generation of a word from a part-word stimulus, do not correlate with one another. Thus, we are of the belief that we have reliable rehearsal and lexical access variables, we are less certain about what the word completion variables are measuring.

In order to reduce the dimensionality of our data, we first carried out principle component factor analyses. Because of the number of participants in our study, we wanted to reduce the number of variables in our factor analysis to ensure that we had reasonable levels of power. Given the pattern of correlations in Table 3, we decided to enter the counting and alphabet variables as a measure of speeded articulation, the articulation and reading measures that have been used by other researchers, word naming as either an item identification or lexical access variable, lexical decision for words and non-word naming as lexical access variables and finally output time. In Table 4 we present the result of a principle components analysis in which the three factors with eigen values greater than one have been extracted and obliquely rotated. Extraction of these factors accounts for a little over 80% of the total variance in these scores. For ease of interpretation, only those loadings of greater than .30 have been included. The first and second factors correlating with each other (r = .32). However, the third factor did not correlate with either of the first two (r = -.11 and -.03, for the first and second factors, respectively).

The first factor extracted seems to reflect lexical memory measures in that the dominant loadings involve the two lexical and word naming measures. The word naming task thus appears to be best conceptualised as a lexical access task rather than a measure of item identification at encoding. The fact that the reading task loads on this variable gives credence to the idea that this task is, in part, a continuous word naming task and that lexical access is a key feature of this task.

The second factor appears to reflect a speed of articulation measure. Again, the fact that the reading task loads on this factor indicates that the reading task is multi-dimensional and that one aspect of this task reflects articulation speed. Given that the variables that load on this factor have traditionally been interpreted in terms

of rehearsal speed, we have labelled this factor a rehearsal speed variable.

The third factor is dominated by the response output measure of speaking time. The counting time variable has a low loading on this factor and although facilitation with numbers might be an important component of digit span, we are interpreting the factor as a more general measure of verbal output time.

Given the above factor structure, we then calculated factor scores and used these factor scores as independent variables in regression analyses to predict span. When lexical, rehearsal and output composites were added to the regression equation simultaneously, the linear combination predicted span accounting for 46% of the variance in span, F (3,43) = 12.19. What is most important is the coefficients of all three variables were statistically significant.

In order to examine the relative importance of the three independent variables, a series of hierarchical regressions were carried out. With factor scores that have been generated from the oblique factors, the order in which independent variables are entered have a bearing upon the amount of variance explained. Table 5 presents three hierarchical regressions where the three variables have been entered in different orders. The first thing to note is that the lexical access variable makes a significant contribution to the variance in span whenever it is entered into the equation; F (1, 45) = 18.21, F (1, 44) = 11.29 and F (1, 43) = 12.18 for analyses 1, 2 and 3, respectively. In contrast, the rehearsal variable makes a reasonable contribution when it is entered before the lexical access variable, F (1, 45) = 7.84 and F (1, 44) = 9.85 for analyses 2 and 3, but adds less if it is added after the lexical access variable, F (1, 44) = 3.41, p = .07. Finally, it is again clear that output time accounts for 12 -15% of the variance whenever it is entered, F (1, 43) = 9.56, F (1, 43) = 9.56 and F (1, 45) = 11.94 for analyses 1, 2 and 3 respectively.

Pause Times, Span and Redintegration.

Cowan (1992; Cowan et al., 1994; Cowan, Wood, Wood, Keller, Nugent & Keller, 1998) have suggested that pause times during verbal output reflect attempts by subjects to retrieve the next item from memory. Cowan's approach is to take lists of a fixed length and look at pause times for these lists. Given the pause times in Table 2, where pause times for six item lists were shorter for those with a maximal span of 7 (mean = .70) than those with a maximal span of 6 (mean = .29), it is reasonable to suppose that pause times should be shorter for the high span group than the group with a maximal span of 7. Presumably, with this measure there should be a correlation between pause times and span, but more importantly, it might be the case that pause times might now correlate with lexical access or rehearsal variables.

To test these ideas we looked first at pause times for six digit lists. In the top panels of Figure 1 we have averaged the pause times in the six digit lists for the low span group (maximal span = 6), the medium span group (maximal span = 7) and the high span group (maximal span = 8 or 9) and plotted these scores against group averages for normal digit span, lexical factor scores, and rehearsal factor scores. In the bottom panels, we have plotted the same variables for seven digit lists and we have decomposed the high level group into two separate groups.

In these figures it is clear that with this method of scoring, pause times now correlate with span. Interestingly, pause times correlate reasonably highly with the lexical access measures and not with rehearsal measures. In short, it seems reasonable to propose that during pauses in verbal output, subjects are accessing lexical memory in order to redintegrate a degraded trace (Browne & Hulme, 1995). The lack of any correlation between pause times and rehearsal measures also supports Cowan et al.'s

(1994) assertion that pause times are not related to rehearsal processes.

Experiment 2

While the results of Experiment 1 appear to be consistent with what is known of immediate serial recall performance, the low number of subjects relative to the number of variables we have used, casts some doubt on the reliability of our results. Secondly, we wanted to see if the way in which we presented some of the lexical tasks had an impact upon performance. In Experiment 2 we again used a number of lexical access tasks, but rather than take measures of individual items, we presented a number of items and measured how many responses could be made in a given time period. In effect these are continuous lexical access tasks. If we can replicate the results of Experiment 1 then it would be a clear demonstration that the effects we observe are quite robust across the different means of assessing performance. In Experiment 2 we thus attempted to replicate Experiment 1 using a larger pool of participants and fewer variables.

Subjects

A total of 130 undergraduate students from the University of Southern Queensland participated in this study. All subjects received course credit in return for their participation. Three subjects produced missing data for one of the measures so consequently the analyses are based upon the performance of the remaining 127 subjects for whom we have complete data. None of the subjects in this experiment participated in Experiment 1.

Measures and Procedure

Subjects were tested individually in a single session in which digit span, counting time, alphabet recitation time, articulation rate, reading rate, lexical decision time, and nonword naming time were all measured. The characteristics of each task were as follows.

Digit span. This task was identical to the one in Experiment 1 save that a greater range of list lengths was used and subjects read the digits silently rather than aloud. The starting length was four items and the length of each list increased by one digit every four trials. With a starting point of four digits and a maximum length of 9 digits, a total of 24 trials were presented to each subject.

Counting, Alphabet Recitation and Articulation Rate. These measures were identical to those used in Experiment 1.

Reading Rate. The fifty low-frequency words that were used in Experiment 1 were again used here. The words appeared on a computer screen in five columns of 10 words. To present this task, the computer screen was blank for a period of five seconds. The five columns then appeared simultaneously on the screen for a period of 15 seconds and then the screen went blank again. While the words were on the screen participants read as many of the words as they could. The dependent measure here is the number of items that were read in that 15-second period. It should be noted here that, unlike Experiment 1, where the dependent measure was the time spent vocalising, the current measure reflects vocalisation time, preparation time and any pauses that were produced while reading.

Nonword Naming. Forty-two words from the Toronto Word Pool (Friendly et al., 1982) were selected and one letter changed to make a pronounceable nonword (e.g. abort – amort). The forty-two nonwords were arranged in three columns and were presented on a computer screen in the same manner as the Reading task. Participants had 15 seconds to pronounce as many of the nonwords as they could. Again the measure contains pause times, preparation time and vocalisation time.

Lexical Decision. The 20 words and 20 nonwords that were used in Experiment 1 were again used. The words were presented simultaneously on the computer screen in randomised order in four columns. Participants were instructed to look at each letter string and decide as quickly as possible whether or not the string was an English word. The letter strings appeared on the screen for 20 seconds and the number of 'yes' and 'no' responses generated in that period served as the dependent measure. Note that in this experiment separate measures for lexical decision to words and non-words have been collapsed into a single measure.

Results

The descriptive statistics for the measures used are presented in Table 6. Direct comparisons between Experiments 1 (Table 1) and 2 (Table 6) can be made for the span, articulation, alphabet recitation and counting tasks. Mean levels of performance on these tasks are very similar, with performance being more variable in Experiment 2. In comparing reading, lexical decision, and non-word naming, we have re-scored the data in a way that performance can be compared. For the reading task, participants in Experiment 1 read at the rate of .51 seconds per word; in Experiment 2 they read at the rate of .53 seconds per word. With non-word naming, it took participants .90 seconds say each non-word in Experiment 1 and .87 seconds per non-word in Experiment 2. With the lexical decision task, separate estimates could be made for words and non-words in Experiment 1 with lexical decision time for words being .85 seconds and 1.12 seconds for nonwords. In Experiment 2, average response time was .76 seconds, collapsed across words and nonwords. In short, there are no great changes in performance across the two Experiments.

The correlations with span again replicate those of Experiment 1. The lexical access variables again correlate with span, as do reading and articulation rate measures. In this experiment, there is a significant correlation between span and counting time, which was not evident in Experiment 1, but has been observed in other data (Cowan et al., 1998).

The outcomes of a principle components factor analysis with an oblique rotation are presented in Table 7. A two-factor solution accounted for 66% of the variance. The two factors in this experiment are almost identical to factors 1 and 2 in Experiment 1, accounting for equivalent amounts of variance. Factor 1 is defined by the lexical access variables and factor 2 is defined by the rehearsal variables. The correlations between the two factors are very similar as well; $r = .32$ in Experiment 1 and $r = -.34$ in Experiment 2 (note that the change from positive to negative correlation reflects the change in the way the lexical access variables were measured). The only change across experiments is that reading loaded on both factors in Experiment 1 whereas it loads solely upon the lexical access variable in this experiment (the loading on the rehearsal variable was -.15).

Hierarchical regressions were carried out in which span was predicted from factor scores derived from the principle components analysis. The lexical and rehearsal variables accounted for approximately 20% of the variance in span, F (2, 124) = 15.09. The results of the hierarchical analyses are presented in Table 8. When lexical variables were first added, 19.6% of the variance was explained, F (1, 125) = 29.67; adding the rehearsal variables increased the explained variance by a non-significant .4%, F (1, 124) = .61. When rehearsal variables were added first, a statistically significant 5% of the variance was explained, F (1, 125) = 5.64. On entering the lexical variables a further 15% of the variance was explained, F (1, 124) = 23.53. As was the case in Experiment 1, lexical access variables seem to make a

more significant contribution to individual differences in span than do rehearsal variables.

Discussion

For the past two decades the main features of performance on immediate serial recall have been well accounted for by the articulatory loop model of Baddeley and his colleagues (Baddeley, 1986; Baddeley et al., 1975) and other similar trace decay models (Schweickert & Boruff, 1986). More recently, there have been suggestions that long-term memory processes make a large contribution to performance (Hulme et al., 1991), as do forgetting and retrieval dynamics during actual recall (Cowan, 1992). In these latter cases, much of the work has been done with children as the subject population. The current research was motivated by an attempt to see if the suggestions made apply equally to an adult sample and to see if hypothesised rehearsal, output and long-term lexical contribution to span performance could be examined within the one experiment.

The results are quite straightforward. We replicate a number of effects that have been observed by other researchers and show that many of these effects are observed with adult samples, not only with children. We show that reading rate and articulation rate are both correlates of span (Baddeley et al., 1975; Hitch et al., 1989; 1993; Schweickert & Boruff, 1986), and that span and word naming time correlate (Case et al., 1982; Hitch et al., 1989; 1993). We show that adult subjects speak at much the same rate when they output their responses for span length lists irrespective of how good at the task they are. Our more exploratory analyses indicate that we can produce reliable composite measures for lexical access, rehearsal speed and output speed variables. In so doing, we show that variables like reading rate speed may not be a pure measure of rehearsal speed, as has often been assumed in past research. Furthermore, we are able to show that lexical access appears to be a more important determinant of span than either rehearsal speed or output time, although the latter do make significant contributions to span. In the following sections we want to examine the impact of these results for the different models and theories that were outlined in the introduction to the current research.

Trace decay models argue that span performance is supported by a rapidly decaying trace that can be refreshed and maintained in a useable form by rehearsal. There are a number of empirical phenomena that have been cited to support these models, but for present purposes there are three that are of interest. Firstly, there is a strong correlation between span and rehearsal rate. The second is that the slope of the regression line for predicting span from rehearsal rate is around 2, (Baddeley et al., 1975; Brown & Hulme, 1995; Standing et al., 1980; Schweickert & Boruff, 1986). Thirdly, that span is equivalent to what can be output in two seconds (Schweickert & Boruff, 1986).

The above features appear to be present in our data set. In both experiments correlations between span and the traditional reading rate and articulation rate measures are reliable. The slopes of the regression equations for predicting digit span from word reading rate were 1.59 in Experiment 1 and 2.07 in Experiment 2. Although, it is more traditional to predict span from reading rate using the same set of materials, the differences here are not sufficiently great to assume that anything odd is occurring in the data. Finally, the longest lists recalled in order are output in 2.4 seconds, on average, which again is within reasonable bounds. Our results appear to replicate the relevant findings that are taken to support trace decay plus rehearsal models. As such, there appears to be little reason to reject these models. However, we

think that there are aspects of the data that do challenge these models.

The first challenge we would like to address involves the rehearsal measure. Measuring covert rehearsal has been a long-standing problem with researchers employing such diverse methodologies as measuring EMG and changes in pupil size, button pressing during rehearsal, and pause times during learning to obtain rehearsal measures (see Johnson 1980 for a review). Alternatively, researchers have tried to avoid measurement problems by getting subjects to rehearse overtly, or to measure some overt behaviour that is thought to correlate with the covert process. The reading and articulation measures that Baddeley et al. (1975) used were examples of the latter approach. We suggested earlier that doubts had been raised as to whether reading rate and articulation rate were good and valid measures of rehearsal. Our analyses indicated that the articulation measure is a good measure of speeded verbalisation in that it clearly aligns with the counting and alphabet recitation tasks. However, the problem remains that while verbalisation speed may be a component of covert rehearsal, it is not at all clear that it is the sole component or that it is the most important component. There are also problems with reading speed as a measure of rehearsal. The factor analyses in Experiment 1 indicated that reading was not a pure measure of rehearsal. That is, it loaded upon both the lexical access factor and the articulation speed factor rather than on the articulation speed factor alone. In Experiment 2, the reading variable loaded on the lexical access factor alone rather than on the two factors. The results clearly indicate that reading time is not a good measure of rehearsal rate, instead it seems to be a measure of speed of lexical access.

The second challenge to the trace decay models involves the assumption that in the absence of rehearsal, the trace lasts for an invariant period with the result that span is the equivalent of what can be output in two seconds. As Cowan (1992) has pointed out, the fixed trace duration model implies that those with high spans should be producing more words per unit of time during output than those with low span. In the current data set, as was the case with Cowan's sample of children, this just does not happen. High span and low span subjects speak at the same rate, with high span subjects talking for longer than those with short spans do. Moreover, approximately, 20% of the participants in Experiment 1 are taking longer than 3 seconds to recall a list in correct order. Our results support several studies indicating that the trace that supports recall can survive for longer than 2 seconds (Dosher & Ma, 1998; Longoni, Richardson & Aiello, 1993; Tolan & Tehan, in press). In short, the current results do not appear to be consistent with the assumptions about fixed trace duration of two seconds.

The current data may also have implications for how one interprets the data that have been used to argue for fixed trace duration. One of the often-quoted examples is the Stigler et al. (1986) data set that compared maximal span of Chinese and American students. While the span of Chinese students was reliably higher than that of the Americans, Chinese digits were pronounced more quickly than their English counterparts. When output was measured, there was no significant difference between output times with output of maximal span being about 2.4 seconds for Chinese students and 2.9 seconds for the American students. We are interested in this data set because the Australian students in the current experiments look very similar to the US college students. For the Australian sample in Experiment 1, average maximal span was 7.10 (SD = 1.09) and for the US sample was 7.16 (SD = 1.04). Output response time (speaking time plus interword pauses) was 2.57 seconds (SD = .62) for the Australian students, and was 2.90 seconds (SD estimated to be in the

region of .69) for the US students. The aspect of the Stigler et al. data that was not emphasised in their paper was the distribution of output times for maximal span length lists. The distributions of Australian and American samples are roughly equivalent, but in the Australian sample, output time is related to individual differences in span. Given the similarities between the two samples, it is possible that those American students who had longer output times also had higher spans. If this speculation is correct, the American sample may well be a better example of the case against the trace decay models than it is for it.

While the results challenge simple trace decay models, they clearly support those models that involve a contribution from long-term lexical memory (Brown & Hulme, 1995; Schweickert, 1993). Hulme and his colleagues have numerous data sets in which span for short and long lists of different types of material have been correlated with rehearsal rates for the same sets of materials. In all the data sets they find a strong linear relationship between span and reading rate. When words or familiar non-words are used as the to-be-remembered material, the slope of the regression line generated for predicting span from reading rate, is parallel for most materials. The differences in span for the different types of material are reflected in the intercept of the regression equation. Hulme has argued that this intercept reflects the contribution of lexical memory to span. The current factor analyses confirm the linkage of lexical decision and naming task as tasks that are primarily measuring speed of lexical access. The regression analyses suggest that facility in lexical access is, as Hulme suggests, an important determinant of span. Moreover, these analyses indicate that it is more important than either rehearsal or output speed in determining individual differences in span.

The fact that lexical decision time is a correlate of span is a novel finding that has a number of implications. As mentioned earlier, there is some dispute as to what measures like word naming are measuring. With lexical decision there is no debate about what is being measured. This has ramifications for explanations of the word frequency effect. Both Tehan and Humphreys (1988) and Gregg, Freedman and Smith (1989) found word frequency effects in immediate serial recall under both articulatory suppression and rehearsal conditions. Gregg et al. argued for a two-component explanation. When subjects were free to rehearse, they argued that the frequency effects could be attributed to the fact that high frequency words were rehearsed faster than low frequency words, whereas under suppression these effects were attributable to recall from long-term memory. Tehan and Humphreys argued that parsimony would suggest a long-term memory contribution in both instances. The current results lend support to the latter explanation and pinpoint the locus of the effect to access of lexical memory (see Hulme et al., 1997 for a similar point).

The other implication of the correlation between span and lexical decision is that there is a rich research heritage associated with the lexical decision task. For example, the task has been widely used to explore various forms of priming effects (repetition, phonological and semantic priming effects), and has served as a vehicle for exploring lexical neighbourhood effects (Andrews, 1989). These areas have largely been ignored in immediate serial recall research. Our research hints that these might be areas that could profitably be explored.

While the regression analyses indicate that speed of lexical access is an important determinant of span, they are silent as to the locus of the effects. For instance, some models clearly assert that efficiency in activating cognitive units from perceptual input resulted in more resources available for storage (Case, et al., 1992;

Dempster, 1981). If one equates activating cognitive units with accessing lexical memory then these models suggest that speed of access to lexical memory is having its effect at encoding. The same argument might be made from the trace decay perspective. The faster that you can identify an item, the more time there is available for rehearsal and the more rehearsal, the better the span[1].

Other models, notably those that involve a redintegration process (Brown & Hulme, 1995; Schweickert, 1993), posit that the primary effects of lexical access are at the retrieval stage. In these models, recall is based upon a degraded trace and the subject's task is to generate an item from this degraded trace. From an individual differences approach, those subjects who are better able to access lexical memory should do a better job of redintegrating degraded traces.

Our own biases are such that we favour the redintegration approach, in that we think that lexical access effects should be enhanced under conditions where such access is more difficult. Here we assume that reading a word at input, where there is little degradation, is not going to produce lexical access problems in the same way that trying to generate a word from a degraded stimulus will. Furthermore, we think that there is evidence in favour of the redintegration approach in the data. The fragment completion task was included in Experiment 1 as an attempt to get at the process of generating an item from a degraded stimulus. The correlations reported in Table 3 indicated that this task shares much in common with the other standard lexical access tasks. More direct evidence comes from the pause times of the six-item and seven-item lists. The correlation between pause times and lexical access factor scores suggests that subjects are accessing lexical memory during these pauses and are presumably trying to reconstruct a degraded memory trace.

One further point needs to be raised with respect to the relationship between rehearsal and item identification measures. The study by Hitch et al. (1993) indicated that rehearsal rate made a more substantial contribution to span than did item identification time. In that experiment, rehearsal rate (articulation time) and item identification (word naming) speed were correlated, but when the variables were entered into a regression equation, rehearsal rate always explained a significant amount of variance, irrespective of when it was entered. Item identification time only made a contribution to span if it was entered first. The Hitch et al. results appear to be the opposite of the findings in Experiment 1 when our composite measures are used as predictor variables. In exploring this discrepancy we looked at the contribution of individual measures of rehearsal speed and the contribution they made with word naming to span.

Of the four rehearsal speed measures we employed in Experiment 1, we first chose the reading rate measure as the predictor variable for rehearsal speed. Regression analyses produced identical results to that obtained by Hitch et al.; the rehearsal rate measure made a significant contribution to span irrespective of when it was entered, but item identification time only contributed when it was entered into the equation first. That is, we replicated Hitch et al.'s demonstration that rehearsal contributes more to span than does item identification speed. However, when we chose articulation time as the measure of rehearsal speed (the measure that best approximates the Hitch et al. procedure) both rehearsal and item identification made roughly similar contributions. Moreover, when counting time was used as the rehearsal measure, rehearsal speed contributed nothing to span whereas item identification did. Clearly, the relative contribution of rehearsal speed and item identification time to span depends upon the measure chosen.

The other aspect of our results that requires comment is the response output features. We replicate Cowan's (1992) findings with 4-year-old subjects; higher spans are reflected in longer output times. As far as pause times go, the impression given in the Cowan paper is that the children paced their recall leaving pauses between words. Cowan suggested that during these pauses, his subjects reactivated the memory trace and searched through it for the next item. If this process is being employed by adults then the activation and search appears to be going on in parallel with the output of previously retrieved words, in that the pauses that seem characteristic of the children's performance are not present to the same extent with the adults. What pauses there are, seem to be brief and towards the end of the list. Having said this, what pauses there are become more pronounced as lists increase in length towards span. The pauses that were present were informative in that is seems that as span approaches, subjects need extra time to reconstruct the degraded memory trace of the last one or two items.

While we have shown that lexical access is a contributor to individual differences in span, there are other facets of immediate memory that we have not measured or not taken into account that might well make a significant contribution to span. Most of the current models of span suggest that recall is based primarily upon phonemic representations of the items. None of the measures that we have included have been directed at this aspect of span performance. Individual differences in the use of transient phonemic information may well be an important part of span performance.

A related issue involves susceptibility to proactive interference. Tehan and Humphreys (1995; 1998) have suggested that immunity to proactive interference at very short retention intervals, is related to the availability of phonemic codes. If there are individual differences in the encoding, maintenance, or retrieval of phonemic codes, then proactive interference might well begin to exert an influence on span.

The span task is essentially a task that involves retrieval of order as well as retrieval of items. In the current research we have concentrated upon factors that primarily affect item information. We have not specifically examined the extent to which maintenance of order information is a powerful determinant of span.

Finally, we have largely ignored the fact that not all subjects perform the span task in the same way. Logie, Della Salla, Laiacona, Chalmers & Wynn (1996) have shown that a wide variety of strategies are normally used to perform the span task. We have made no attempt to disambiguate strategy use in the current experiments. It remains possible that the mix of rehearsal, lexical and output processes might differ for someone who uses a verbal rehearsal strategy as opposed to someone who uses chunking or someone who adopts a visual imagery or semantic approach to the task. However, while the mix of processes might change, we would still want to assert that our data indicate that a number of different processes are involved in immediate recall of short lists.

Conclusion

For the last two decades, trace decay models have tended to dominate our understanding of memory span performance. The idea that speed of rehearsal is a prime determinant of span has gone unchallenged until relatively recently. This is not surprising in that influences on span such as word length and articulatory suppression are easily understood from a trace decay perspective. In recent times, however, there has been a shift in emphasis towards other processes being involved in span. Dosher and Ma (1998) have indicated that forgetting processes during output are better predictors of span than rehearsal measures. Cowan et al. (1998) suggest that retrieval

search processes during output are at least as important as rehearsal. The current research also confirms that output processes are at least as important as rehearsal. However, we make a further contribution by showing that lexical access is a key determinant of individual differences in span, much more so than either rehearsal or output time. The other feature of the data is that some of the tasks that have been used in the past to measure rehearsal speed may not be pure measures of rehearsal speed. What is clear is that simple speed of articulation does not have much of an impact upon individual differences in span. Our research suggests that there is a need for better measures of rehearsal, if not for better descriptions of what rehearsal is.

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search. Journal of Experimental Psychology: Learning Memory and Cognition, 15, 802-814.

Atkinson, R.C. & Shiffrin, R.M. (1968). "Human memory: a proposed system and its control processes". In K.W. Spence and J.T. Spence (eds.) The Psychology of Learning and Motivation, (Vol 2., pp. 89-195). New York: Academic Press.

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.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monographs, 80(3, Pt. 2).

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

Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A lexical contribution to serial order recall of subspan lists. Psychonomic Bulletin and Review, 1, 122-125.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33, 386-404.

Chase, W. G., Lyon, D. R. & Ericsson, K . A. (1981). Individual differences in memory span. In M.P. Friedman, J . P. Das, and N. O'Connor (Eds.) Intelligence and Learning. (pp. 157-162.) New York: Penlum 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.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. British Journal of Psychology, 55, 429-432.

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

Cowan, N. (1993). Activation, attention and short-term memory. Memory and Cogntion, 21, 162-167.

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.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. Psychological Bulletin, 89, 63-100.

Dosher, B. A. & Ma, J. J. (1998). Output loss or rehearsal loop? Output time versus pronunciation time limits in immediate recall for forgetting matched materials. Journal of Experimental Psychology: Learning Memory and Cognition, 24, 316-335.

Elliott, J. M. (1992). Forward digit span and articulation speed for Malay, English, and two Chinese dialects. Perceptual and Motor Skills, 74, 291-295.

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. British Journal of Psychology, 71, 43-51.

Estes, W.K. (1972). An associative basis for coding and organization in memory. In A.W. Melton and E. Martin (Eds.) Coding Processes in Human Memory. (pp. 161-190.) Washington, D.C.:Winston.

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.

Gregg,V. H., Freedman, C. M. & Smith, D. K. (1989). Word frequency, articulatory suppression and memory span. British Journal of Psychology, 80, 363-374.

Henry, L. A., & Millar, S. (1991). Memory span increase with age: A test of two hypotheses. Journal of Experimental Child Psychology, 51, 459-484.

Henson, R. N. A., Norris, D. G., Page, M. P. A. & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. Quarterly Journal of Experimental Psychology, 49A, 80-115.

Hitch, G. J., Halliday, M. S., & Littler, J. E. (1989). Item identification time and rehearsal rate as predictors of memory span in children. The Quarterly Journal of Experimental Psychology, 41, 321-337.

Hitch, G. J., Halliday, M. S., & Littler, J. E. (1993). Development of memory span for spoken words: The role of rehearsal and item identification processes. British Journal of Developmental Psychology, 11, 159-169.

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., 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., Roodenrys, S., Schweickert, R., Brown, G.D.A, Martin, S. & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. Journal of Experimental Psychology: Learning Memory and Cognition, 23, 1217-1232.

Johnson, R. E. (1980) Memory-Based Rehearsal. In G. Bower (Ed.) The Psychology of Learning and Motivation, (Vol 14, pp 263-307.) New York: Academic Press.

Kucera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence. RI: Brown University Press.

Logie, R. H., Della Salla, S., Laiacona, M., Chalmers, P. & Wynn, V. (1996).

Group aggregates and individual reliability: The case of verbal short-term memory. Memory and Cognition, 24, 305-321.

Longoni, A. M., Richardson, J.T.E. & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. Memory and Cognition, 21, 11-22.

Naveh-Benjamin, M., & Ayres, T.J. (1986). Digit span, reading rate, and linguistic relativity. The Quarterly Journal of Experimental Psychology, 38A, 739-751.

Nelson, D.L. & McEvoy, C.L. (1979). Encoding context and set size. Journal of Experimental Psychology: Human Learning and Memory, 5, 292-314.

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of lexical factors in immediate serial recall. Quarterly Journal of Experimental Psychology, 48A, 384-404.

Peterson, L.R., & Johnson, S.K. (1971). Some effects of minimizing articulation on short-term retention. Journal of Verbal Learning and Verbal Behavior, 10, 346-354.

Roodenrys, S., Hulme, C., Alban., Ellis, A. W., & Brown, G. D. A. (1994). Effects of word frequency and age of acquisition on short-term memory span. Memory and Cognition, 22, 695-701.

Salamé , P., & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. Journal of Verbal Learning and Verbal Behavior, 21, 150-164.

Schweickert, R. S. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. Memory and Cognition, 21, 168-175.

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.

Schweickert, R., Guenert, L & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. Psychological Science, 1, 74-77.

Standing, L., Bond, B., Smith, P., & Isely, C. (1980). Is the immediate memory span determined by subvocalization rate? British Journal of Psychology, 71, 525-539.

Stigler, J. W., Lee, S., & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. Cognition, 23, 1-20.

Tehan, G., & Humphreys, M. S. (1988). Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note. Bulletin of the Psychonomic 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. & Humphreys, M. S. (1998). Creating proactive interference in immediate recall: Building a dog from a dart, a mop and a fig. Memory and Cognition, 26, 477-489.

Tolan, G. A. & Tehan, G. (in press). Determinants of short-term forgetting: Decay, retroactive interference or proactive interference? International Journal of Psychology.

Table 1
Descriptive Statistics for all Variables, and Correlations with Digit Span in Experiment 1

Measure	M	SD	Correlation with span
Digit Span	6.14	.85	1.00
Preparation Time	1.24	.55	-.05
Pause Time	.91	1.33	-.20
Spoken Time	2.42	.62	.36*
Reading Time	25.43	3.66	-.50*
Articulation Time	11.72	1.37	-.47*
Counting Time	7.79	1.37	-.12
Alphabet Naming	4.45	.95	-.25
Fragment Completion	6.16	2.01	.44*
Ending Completion	4.48	1.84	-.29*
Category Instance Generation	3.17	.99	-.29*
Lexical Decision Word	.85	.18	-.49*
Lexical Decision Non-word	1.12	.31	-.57*
Word Naming	.62	.13	-.40*
Non-word Naming	.90	.24	-.41*

Table 2

Timing Characteristics (secs) of Verbal Output of Span and Subspan-length Lists

	Low (6) N=15	Medium (7) N=15	High(8-9) N=17
Speaking Time			
Span (n)	1.93	2.42	2.87
Subspan (n-1)	-	2.03	2.42
Average Word Duration			
Span (n)	.32	.34	.36
Subspan (n-1)	-	.34	.34
Preparation Time			
Span (n)	1.30	1.16	1.23
Subspan (n-1)	-	1.51	1.23
Pause Time			
Span (n)	.70	.46	.57
Subspan (n-1)	-	.29	.17

Table 3.
Correlations Among Non-span Measures

	ST	RT	AT	CT	AN	FC	CI	EC	LW	LN	WN
Spoken Time (n)	1.00										
Reading Time	-.12										
Articulation Time	-.07	.53									
Counting Time	.29	.35	.59								
Alphabet Naming	.01	.43	.44	.43							
Fragment Comp	-.09	-.43	-.33	-.15	-.03						
Category Instance	-.14	.33	.25	-.01	.11	-.42					
Ending Completion	-.15	.33	.31	-.07	.17	-.14	.39				
LexDec - Word	-.06	.57	.31	.09	.15	-.52	.41	.43			
LexDec - Nonword	-.19	.46	.34	.08	.34	-.54	.37	.43	.74		
Word Naming	-.08	.63	.30	.21	.27	-.40	.32	.36	.83	.73	
Nonword Naming	-.02	.69	.23	.13	.20	-.52	.27	.25	.76	.69	.84

Table 4

Rotated Factor Loadings of Non-span Measures in Experiment 1

	Factor		
	1	2	3
Lexical Decision Word	.93		
Word Naming	.94		
Nonword Naming	.96		
Reading	.59	.43	
Articulation		.85	
Counting		.82	.36
Alphabet		.77	
Spoken time (n)			.98
Eigen Value	3.70	1.66	1.06
% of Total Variance Explained	46.31	20.76	13.29

Table 5.
Hierarchical Regressions: Predicting Span from Lexical, Rehearsal and Output Measures in Experiment 1.

	Cumulative R^2	Change in R^2
Step 1 Lexical	.29	.29*
Step 2 Rehearsal	.34	.05
Step 3 Output	.46	.12*
Step 1 Reheasal	.15	.15*
Step 2 Lexical	.34	.19*
Step 3 Output	.46	.12*
Step 1 Output	.14	.15*
Step 2 Rehearsal	.28	.16*
Step 3 Lexical	.48	.15*

Note: * indicates that the change in R^2 is statistically significant.

Table 6
Descriptive Statistics for all Variables, and Correlations with Digit Span in Experiment 2 (N = 127)

Measure	M	SD	Correlation with span
Digit Span	6.21	.96	1.00
Reading Time	29.48	5.82	.36**
Articulation Time	11.19	1.98	-.23*
Counting Time	7.40	1.87	-.30**
Alphabet Naming	5.21	1.33	-.09
Lexical Decision	15.29	4.59	.28**
Non-word Naming	17.25	5.86	.45**

Table 7

Rotated Factor Loadings of Non-span Measures in Experiment 2

	Factor	
	1	2
Lexical Decision	.82	
Nonword Naming	.83	
Reading	.79	
Articulation		.60
Counting		.73
Alphabet		.94
Eigen Value	2.76	1.19
% of Total Variance Explained	46.09	19.94

Table 8
Hierarchical Regressions: Predicting Span from Lexical Access and Rehearsal Measures in Experiment 2.

	Cumulative R^2	Change in R^2
Step 1 Lexical	.192	.192*
Step 2 Rehearsal	.196	.004
Step 1 Reheasal	.043	.043*
Step 2 Lexical	.196	.153*

Note: * indicates that the change in R^2 is statistically significant.

Author Notes

We would like to thank Charles Hulme, Nelson Cowan, Gerry Fogarty and Robert Logie for helpful comments on an earlier draft of this article and Liam Hendry, Craig Smith, Tricia Greening and Kath Ryan for assistance in data collection.
Correspondence concerning this article should be addressed to Gerry Tehan, Department of Psychology, University of Southern Queensland, Toowoomba. 4350. Australia. e-mail tehan@usq.edu.au

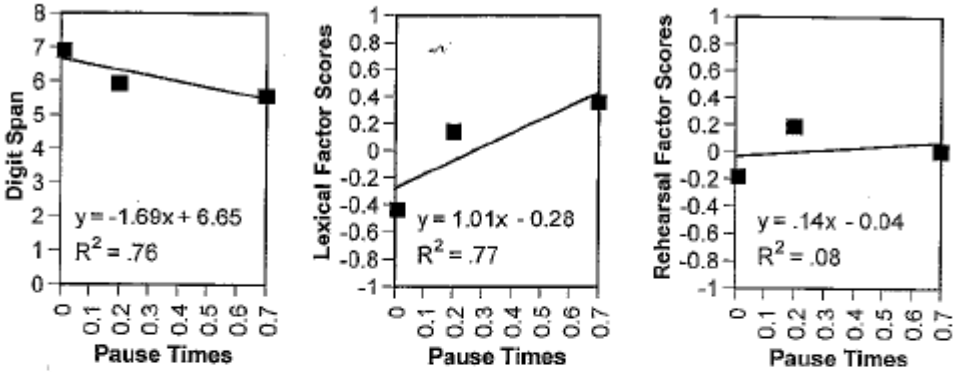
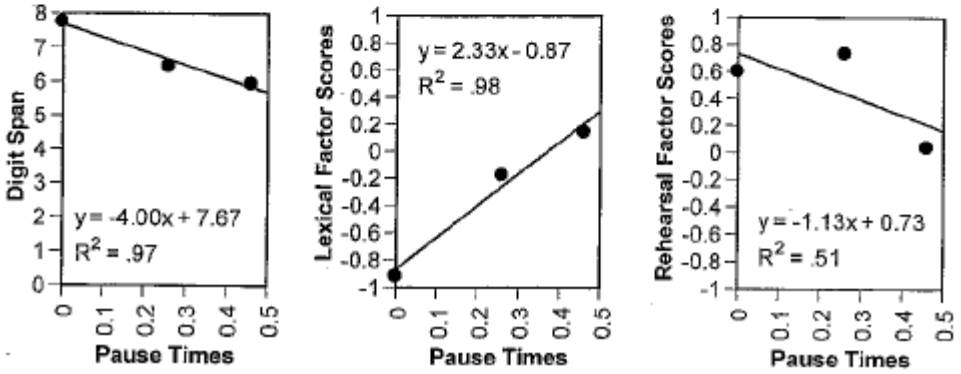
Footnotes

1. We would like to thank Mike Page for pointing out this possibility.

Figure Caption

Figure 1. Digit span, lexical access speed, and rehearsal speed as a function of pause times on six item lists (upper panels) for high span (left – max span = 8 or 9), medium span (middle – max span = 7) and low span (right – max span = 6) groups and pause time on seven item lists (lower panels) for high span (left – max span = 9), medium span (middle – max span = 8) and low span (right – max span = 7) groups.

Pauses in Six Digit Lists



Pauses in Seven Digit Lists