

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, and KATHERINE RYAN
University of Southern Queensland, Toowoomba, Queensland, Australia

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. In the present research, we used structural equation modeling to explore, at the construct level, individual differences in immediate serial recall with respect to rehearsal, search, phonological coding, and speed of access to lexical memory. We replicated standard short-term phenomena; we showed that the variables that influence children's span performance influence adult performance in the same way; and we showed that speed of access to lexical memory and facility with phonological codes appear to be more potent sources of individual differences in immediate memory than is either rehearsal speed or search factors.

Most current models of immediate serial recall (ISR) have three basic assumptions. The first is that short-term information is lost very rapidly and, as a result, serial recall is often based on a degraded memory trace. Some models attribute forgetting to decay (Baddeley, 1986; Burgess & Hitch, 1999; Cowan, 1999; Page & Norris, 1998); 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 reintegrated in some manner (Brown & Hulme, 1995; Farrell & Lewandowsky, 2002; Schweickert, 1993). The third assumption is that the trace that supports recall is speech based in nature. That is, immediate recall relies on phonological codes. Although 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 ISR. This approach is adopted in the present study.

The standard account of immediate memory (Nairne, 2002) takes the view that performance is the outcome of a balancing act between rapid deactivation of the memory trace and rehearsal, where rehearsal is assumed to refresh and reactivate 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 3 words as quickly as possible 10 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 sec. 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 have been 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 in which 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,

The authors thank Robert Greene, Elisabet Service, and two other reviewers for very helpful comments on an earlier version of this article. Correspondence regarding this article can be addressed to either G. Tehan or G. Fogarty, Department of Psychology, University of Southern Queensland, Toowoomba, QLD, 4350 Australia (e-mail: tehan@usq.edu.au or fogarty@usq.edu.au).

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 et al., 1992) presented an alternative account of the rapid forgetting in immediate memory that is based on decay during the recall process. They demonstrated that longer words at the start of a list caused more forgetting for the words at the end of a list than occurred when 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 output with the articulation of long words than with that of 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 nonrecalled items would be minimized.

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 increased in difficulty, participants produced longer and longer pauses between list items. Given a memory architecture in which 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 et al. (1998) tested these ideas, using an individual differences approach, and were able to show via structural equation modeling (SEM) that both articulation and search measures 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 five-syllable words or for one- and five-syllable nonwords. As such, the role of memory search rate is worthy of further exploration.

It is now commonly accepted that a critical process in ISR is the ability to reconstruct or reintegrate a degraded short-term trace. Although current models of immediate memory incorporate such a stage (Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Page & Norris, 1998), these models are not overly specific about the memory system that is accessed during this process. It is becoming increasingly clear that reintegration 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 nonwords (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, and 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 (Brown & Hulme, 1995; Hulme et al., 1991; Schweickert, 1993).

The assumption that lexical access involves phonemic traces is based on the finding that immediate recall seems to rely primarily on 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. (1991) 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, Tehan and Lalor's (2000) research 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 was pointed out in the introduction to this article, 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) emphasized the role of phonological coding.

Although 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

(2000) 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 aims above 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.

Rehearsal. In the original work exploring the relationship between span and rehearsal rate (Baddeley et al., 1975), two indirect measures of rehearsal were utilized: reading a list of 50 words as quickly as possible and repeating 3 words as quickly as possible 10 times. The reading time estimate of rehearsal speed has faded from use, due to the complexity of the reading task. Moreover, in Tehan and Lalor's (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 experiments in which rehearsal of three-, four-, or five-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 artifact 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, and Tilley (2002) explored speech rate in children for one-word and three-word lists. Like Hulme et al. (1984), they found that both measures correlated with span, with the correlation being 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 research above 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 Tehan and Lalor (2000). 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 backward from 20 to 1 as quickly as possible. The 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. Although some might wonder about the inclusion of nonword naming as a measure of lexical access, given that by definition nonwords have no lexical representations, nonword naming has been important to the development of *dual-route* models of lexical access (Coltheart, 1980).

Tehan and Lalor (2000) 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 present 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, such as the extent of the development of lexical memory, the age of acquisition, or neighborhood effects, may have an effect upon span but are not the focus of the present 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 is quite a range of measures of phonological ability. Some of the more common tasks involve the following: phoneme deletion ("what word results if the first sound of the word *spark* is deleted?"); nonword repetition (say the word "*bliborg*"); phoneme blending, where the ability to blend a sequence of sounds into nonwords is measured; nonword segmentation, where children are asked to identify the component phonemes in a nonword; and sound categorization, which measures the child's ability to detect rhyme and/or alliteration ("which of the following words does not fit: *cat, hat, man, fat*").

Although these tasks all deal with phonological aspects of words, not all 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 nonword repetition loaded on a separate factor in the sound categorization 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 whether there were tasks similar to those reported 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 present study.

The sound-grouping test is almost identical to the sound categorization 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 nonrhyming word. Given the similarity between the sound-grouping task for adults and the sound categorization 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 categorization 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 pseudo-homophone-naming task in which participants are presented with a nonword 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 a measure of phonological processing, in that participants must convert phonological information into lexical information.

Search. As was indicated earlier, Cowan et al. (1998) found that span was related to memory search rate in a Sternberg (1966) type task in which search rate was determined by the slope of the reaction time set size function. Although 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 criticized on the basis of contamination via a memory component, this argument would apply 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 et al. (1999) and the Cowan et al. (1998) studies, there were times at which the participants are holding four or five items in memory. For example, in the Cowan et al. (1998) task, the participants were presented with a matrix of digits or letters. They were then given one, three, 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. (1998) 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, it becomes difficult to provide a more direct measure of memory search in which

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 adopted in the present study. That is, we used 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 the 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, the participants were required to recall short lists of words. Furthermore, we used a large pool of words as the stimulus material, so that each word appeared 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 provides only one source of information. More important, however, most current models of immediate recall are not models of span. They are models of list recall where errors occur. That is, not only are they concerned with recalling items in their correct position, but they are equally focused upon explaining the various patterns of errors that are found in ISR. These models are thus implicitly based on 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 scores for correct recall of position for 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 (2000) used exploratory factor analysis followed by hierarchical regression analysis in both the studies they reported, 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 latent memory span 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. On the basis of the literature, additional contributions were expected from search and phonological-coding processes.

METHOD

Participants

There were 126 participants, consisting of 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.* The 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.* The participants were instructed to count backward from 20 to 1 as quickly as they could. A stopwatch was used to measure the number of seconds taken for the task.

3. *Alphabet recitation.* This task involved speeded verbalization of the alphabet from A to 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 nonword-naming task, and a word-naming task.

4. *Lexical decision.* The participants were presented with a list of 40 low-frequency, five-letter strings from the Toronto Word Pool norms (Friendly, Franklin, Hoffman, & Rubin, 1982). The letters in 20 of these words were substituted to create phonotactically legal nonwords (e.g., abort → amort). The newly created nonwords were then randomly interspersed among the remaining words. The list was divided into four columns and was displayed on the computer screen in its entirety for 15 sec. The 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. Instructions in this and subsequent speed tasks stressed the need for accuracy. The participants were to go as fast as they could, provided that they made accurate responses. Lexical decision performance was calculated by measuring the number of correct decisions within the 15-sec 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. The participants were presented with three columns of 14 nonwords and were instructed to read down each column as quickly as possible, pronouncing each nonword aloud. The nonword list disappeared from the computer screen after 15 sec. Nonword-naming scores were calculated by summing the number of nonwords pronounced within the 15-sec time limit.

6. *Word naming.* The materials for the task included 50 low-frequency five-letter words from the Toronto Word Pool norms. The word list was divided into five columns, and the 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-sec 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 present study.

7. *Sound grouping.* The Sound-Grouping Test (Thurstone, 1936) 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 test used here consisted of 30 such lists. We used the lexical access literature to find a list of 30 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 randomized the sequence of 30 lists. All the participants studied the same list. The participants were asked to quickly circle the word, in each row of the sheet provided, that sounded different from 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. *Pseudowords.* The Turse Phonetic Association Test (Turse, 1940) is basically a pseudohomophone-naming task (nonwords that 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 three columns on the computer screen. The participants were instructed that they would be presented with a series of nonwords that, when pronounced, sounded like legitimate English words. Their task was to read aloud down each column as quickly as possible. Scores on the pseudowords task were calculated by summing the number of pseudohomophones correctly pronounced within the 15-sec 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). The participants searched for English words that were embedded in a 15 × 15 letter matrix. The letters of a word were presented either forward or backward within a single row, a single column, or a diagonal. The list of target words was presented underneath the matrix. The 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-sec period.

10. *Digit search.* This task consisted of an 11 × 12 matrix of the digits 1 to 20. The 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-sec period.

11. *Immediate serial recall.* The task was initially set up to examine proactive interference effects, using a methodology developed by Tehan and Humphreys (1995). In the present version of the task, the 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 were told that they would not know in advance whether it would be a one- 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 1 sec after the fifth word in the first block. This was the signal that the trial was a two-block trial and that the participants were to forget the first block and concentrate on remembering the second block. A row of question marks (???) appeared on the screen di-

rectly after the final word in each list, to indicate that the 7-sec recall period had commenced. The 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.

The participants studied 20 trials, consisting of an equal number of one- 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 et al., 1994). Phonological similarity was manipulated by having half of the one-block and half of the two-block trials consist of items all of which 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 20 trials were randomly ordered, and this order was constant across participants.

Initial examination of the data tended to show that the participants paid more attention to the second block than to the first block, in that there were recall differences between one- and two-block trials, with an advantage for the two-block trials. Although proactive interference (PI) effects were included in the design, we decided not to explore these issues in this article (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 pseudowords. Timing measures were obtained using a standard digital stopwatch, and all verbal responses were recorded by the experimenter. Total testing time was approximately 30 min.

RESULTS

Prior to the main analyses, the 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.

Preliminary Analyses: Bench Mark Results

The data shown in Table 1 can be used to compare findings in the present study with those reported by earlier researchers. Looking at the bottom two rows in Table 1, it can be seen that the 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 important, the present 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 Baddeley et al.'s (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 were correlated with recall. They also support Tehan and Lalor's 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 generalize from span measures to serial recall of fixed length lists. It is clear that the serial recall of fixed length lists used by us was functionally equivalent to a span measure. Finally, as was predicted, the phonological-coding measures were correlated with serial recall performance.

However, there were some instances in which 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 have shown a correlation with span. Inspection of the other relations involving the search measures suggests that they did not behave 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-sec interval allowed for both tasks did not

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

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11
1. Articulation time	2.49	.16											
2. Counting time	1.98	.25	.44										
3. Alphabet recitation	1.62	.24	.48	.60									
4. Lexical decision	15.27	4.59	-.23	-.27	-.08								
5. Nonword 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. Pseudowords	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 similarity	12.28	5.27	-.24	-.23	-.10	.38	.34	.30	-.41	.30	.16	.12	
12. Span dissimilarity	13.96	6.08	-.25	-.25	-.15	.35	.34	.33	-.44	.34	.08	.16	.74

Note—Correlations above .17 were significant at the .05 level.

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. The 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 do the other types of measures 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 SEM. These analyses are presented next.

The Amos 4.01 statistical package (Arbuckle, 1999) was used on the covariance matrix generated by the 12 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) two-index presentation strategy was followed. With this strategy, the standardized 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 a good fit of the model to the data.

In accord with the design of the study, the first model, depicted in Figure 1, described the following: Articulation time, counting time, and alphabet recitation were markers for a rehearsal factor; lexical decision, nonword naming, and word naming were markers for a speed of lexical access factor; sound grouping and pseudowords were markers for a phonological coding factor; and word search, along with digit search, defined a search factor. Recall of phonologically similar and dissimilar words formed two indicators of the ISR factor. In the structural part of the model, the first four factors were treated as correlated predictors of the ISR factor. The model explained 38% of the variance in ISR, and fit statistics were very good (SRMR = .06; CFI = .99). 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, lexical ac-

cess, and rehearsal to ISR. A large correlation ($r = .91$) between the lexical access and the phonological-coding factors suggested that either these two factors could 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 was expected, fit statistics were very good for both models (SRMR = .05; CFI = .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.

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 ISR than are rehearsal and search processes. This is not to say that these other two processes are not involved in ISR. The technique employed in SEM is similar to standard regression, wherein the beta weights can be interpreted as reflecting the *unique* contribution of the different constructs to the construct of ISR. To determine whether other variables are related, we need to go back to the table of correlations. Table 1 shows that all the 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 although 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^2 decreases by just 4%. If lexical access and phonological coding are removed and rehearsal and search are left in the model, R^2 decreases to 11%. The most parsimonious model of all is one that includes just the phonological-coding factor, which, on its own, explains 34% of the variance in ISR.

DISCUSSION

The results of the present study are relatively clear. Articulation speed and speed of search, at least in the way we have measured them, make little contribution to individual differences in ISR. When speed of lexical access is introduced, 24% of the variance is accounted for. This result replicates a prior study by Tehan and Lalor (2000) that likewise showed that speed of lexical access was a more important factor in producing individual differences in span than was articulation speed. However, when measures of facility with phonological coding are

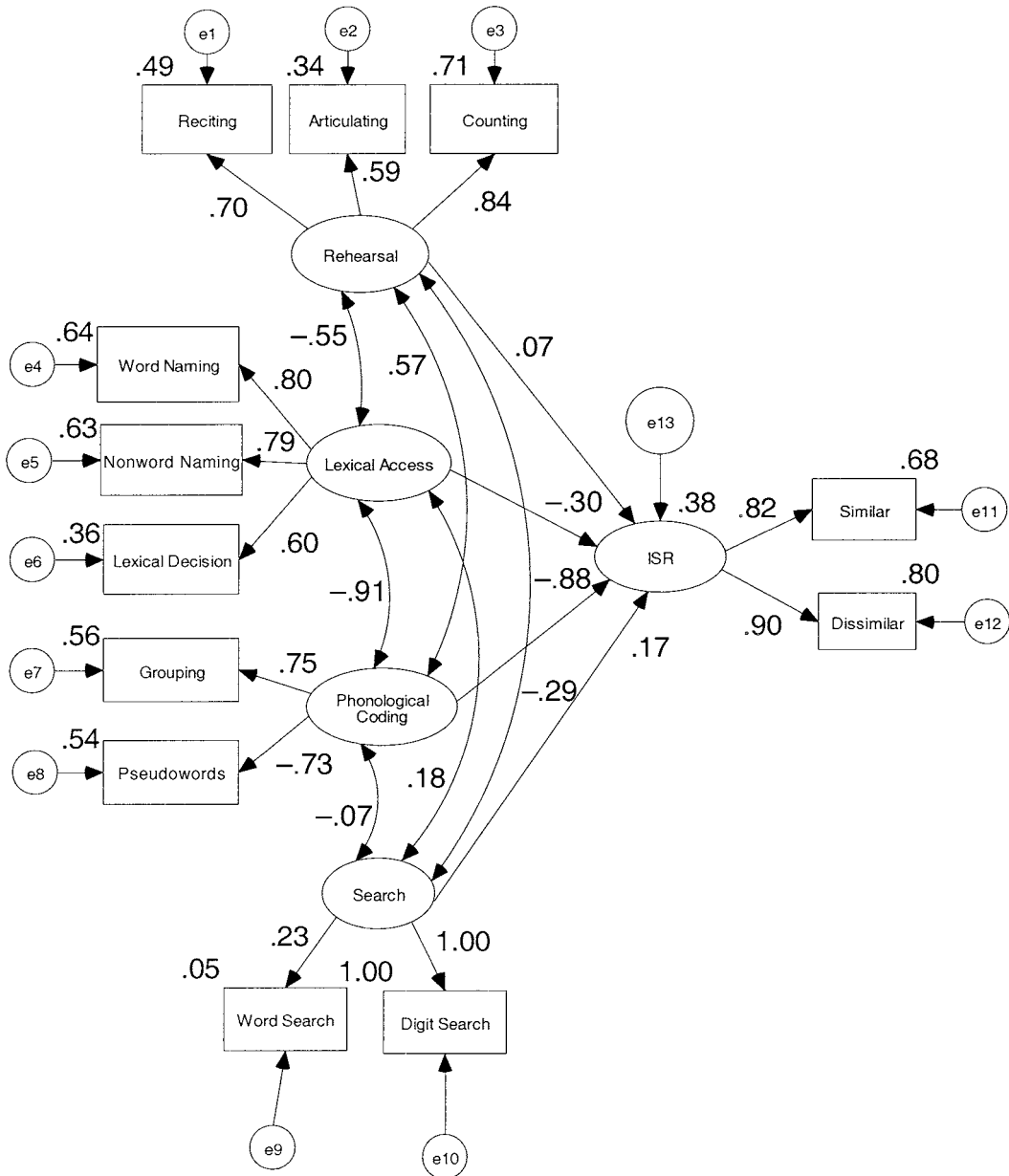


Figure 1. Measurement and structural model for predicting immediate serial recall.

introduced, explained variance rises to 38%, and there are direct linkages from this construct to the ISR construct and to the speed of lexical access construct. In short, phonological coding has a direct influence upon ISR and an indirect influence through speed of lexical access.

In sum, our results indicate that persons who have superior abilities with phonological coding have two advantages. First, they lay down superior traces. Second, if those traces are degraded in any way, they are more skilled at recovering those items from memory.

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, we have

measured search differently from Cowan et al. (1998). Although there is surface similarity between the tasks (both involve search for letters in a large matrix), the derived search measures are quite different. In the present 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. (1998) used a more complex measure, the slope of a regression line. The lack of a correlation with our simple measure may be attributable to the different way in which we have operationalized the search construct.

The Cowan et al. (1998) and the Hulme et al. (1999) 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 in which there is high memory load. 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 minimize memory load.

A second possible 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 present 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 only for approximately 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. The data do not support such an assertion. Speed of articulation contributes only minimal variance to immediate memory performance.

The working memory model and other models of immediate recall assume that items in memory are represented in terms of their phonological characteristics. The fact that there is a direct linkage between phonological coding and ISR is thus not overly surprising, given the privileged role of phonological codes. Our results confirm this basic assumption and suggest that individual differences in the ability to use phonological codes are an important contributor to individual differences in immediate recall. These results reflect similar findings in the developmental literature (Bowey, Cain, & Ryan, 1992; Wagner et al., 1997).

The role of phonological coding via lexical access has less of a history, but such an assumption is a key proposition in Brown and Hulme's (1995) ideas concerning redintegration. That is, at recall, people use phonological information they have available to access lexical memory to try and reconstruct or select an item for recall. Our results are consistent with such assumptions. However, no current accounts of redintegration are developed to the extent that the structure and functioning of lexical memory are incorporated. Word versus nonword differences are readily explained by most models,

but word frequency, word class, and even sublexical 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 realized.

Cowan et al.'s (1998) idea that speed of search through a short-term store is an important determinant of immediate memory is not supported by the data. However, as we have suggested earlier, our measures may not be an adequate operationalization of the search process.

In Cowan's (1999) model, speed of search is only one component. He argues that, during pauses in recall, fading traces are updated and reactivated so that search can continue. If we assume that reactivation is much the same as what others label *redintegration*, our data suggest 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

Although the results may fit well with current ideas concerning the function of phonological representations, there are some caveats. For example, the high correlation between phonological-coding and lexical access measures may simply reflect the fact that neither set of tasks represents a pure measure of the construct. For example, in the pseudohomophone-naming task, participants are asked to use grapheme-to-phoneme conversion rules, but they are required then to access the lexicon to produce an English word. Likewise, in the lexical decision tasks, one possible route that participants may take is, again, to convert graphemes to phonemes and then consult the lexicon. In short, the tasks are roughly equivalent in processes but emphasize different aspects of those processes. One argument against this notion is that the contributions to immediate recall is not equivalent for the two constructs.

The aim of the present research was to test ideas about how immediate recall works. These ideas were the basis for the directionality (and the implied IVs and DVs) among constructs in our theoretical model. Moreover, in light of these ideas, we interpreted the present 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 correlation designs involving latent variables, objections can be raised regarding, first, the identity of the latent variables and, second, the direction of causality. We have argued that lexical decision, word naming, and nonword 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 has pointed out, these tasks are primarily visual-word–

decoding tasks, and what we may well be measuring is the participant's facility in decoding visually presented letter strings. As such, the latent construct might be considered to be speed of item identification at encoding. Tehan and Lalor (2000) considered this possibility and noted that item identification could not be ruled out in their data, as it cannot be ruled out in the present data. However, they noted that lexical access speed correlated with pause times during recall. They argued, as did Cowan (Cowan et al. 1998), that participants were accessing lexical memory during those pauses, presumably trying to reintegrate 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 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, Torgesen, & Rashotte, 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. The present results indicate that the same relationships between measures are present in an adult population (Bowey et al., 1992; Wagner et al., 1994; Wagner et al., 1997). As such, an explanation that phonological coding is important to both span and word decoding is a viable alternative.

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. Although 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 ex-

perimental cognitive psychology, research of the type demonstrated in the present study is, therefore, important for identifying processes that underpin broader cognitive constructs. 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 have noted, previous research emphasizing the role of such constructs as general intelligence and reasoning has tended to overlook the role of contributing processes (Ackerman, Beier, & Boyle, 2002). The fact that the present set of theoretically derived measures was able to account for between 32% and 38% of performance in serial recall indicates the benefits that can be gained from cross-fertilization between the experimental and the individual differences fields.

Conclusion

Our results confirm current thinking that phonological codes have a privileged role in ISR. They have a direct role and an indirect role. The indirect role is assumed to involve a reintegration process in which lexical access from phonological codes is attempted. Individual differences in such abilities produce corresponding individual differences in ISR. We have found much less of a role for other supposedly crucial processes, such as 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: 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 & 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 & 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. S. (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.
- CHASE, W. G., LYON, D. R., & ERICSSON, K. A. (1981). Individual differences in memory span. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 157-162). New York: Plenum.
- 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* (pp. 197-226). London: Routledge & Kegan Paul.
- COWAN, N. (1999). An embedded-process model of working memory. In A. Miyake & 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 & Language*, **31**, 1-17.
- COWAN, N., KELLER, T. A., HULME, C., ROODENRYS, S., MCDUGALL, S., & RACK, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory & 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. E., HOFFMAN, D., & RUBIN, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, **14**, 375-399.
- GATHERCOLE, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & 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 Modeling*, **6**, 1-55.
- 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 & Language*, **30**, 685-701.
- HULME, C., NEWTON, P., COWAN, N., STUART, G., & BROWN, G. D. A. (1999). Think before you speak: Pauses, memory search, and trace reintegration processes in verbal memory span. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 447-463.
- HULME, C., & ROODENRYS, S. (1995). Practitioner review: Verbal working memory development and its disorders. *Journal of Child Psychology & Psychiatry*, **36**, 373-398.
- 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.
- 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 & 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 reintegration in immediate recall. *Memory & Cognition*, **21**, 168-175.
- SCHWEICKERT, R., & BORUFF, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, & 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 Psychonomic Society*, **26**, 293-296.
- TEHAN, G., & HUMPHREYS, M. S. (1995). Transient phonemic codes and immunity to proactive interference. *Memory & 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. *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.
- 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 as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, **33**, 468-479.
- WALLING, J. R., MCEVOY, C. L., OTH, J. E., & NELSON, D. L. (1994). *The University of South Florida rhyme category norms*. Unpublished manuscript, University of South Florida.

(Manuscript received July 2, 2003;

revision accepted for publication December 12, 2003.)