Two Separate Verbal Processing Rates Contributing to Short-Term Memory Span

Nelson Cowan, Noelle L. Wood, Phillip K. Wood, Timothy A. Keller, Lara D. Nugent, and Connie V. Keller
University of Missouri—Columbia

Previous research indicates that verbal memory span, the number of words people can remember and immediately repeat, is related to the fastest rate at which they can pronounce the words. This relation, in turn, has been attributed to a general or global rate of information processing that differs among individuals and changes with age. However, the experiments described in this article showed that the rates of 2 processes (rapid articulation and the retrieval of words from short-term memory) are related to memory span but not to each other. Memory span depends on a profile of processing rates in the brain, not only a global rate. Moreover, there appears to be only a partial overlap between the rate variables that change with age and those that differ among individuals.

An important, current issue in the field of cognition is the extent to which individual differences in mental abilities are uniform across various skills, as opposed to being domain- or task-specific. An issue in the related field of cognitive development is the extent to which observed differences among children can be attributed to differences in maturation, as opposed to reflecting individual variation unrelated to age. Research on response times can be relevant to both of these questions, and was examined here to explore the sources of variation in performance on verbal memory-span tasks.

Recently, tentative answers to these cognitive and developmental questions have been offered. Specifically, some investigators have attributed both individual and developmental differences in human cognition largely to differences in the general or "global" rate of information processing in the human brain. There are several bases for this attribution. IQ is related to the latency of components in event-related brain potentials (McGarry-Roberts, Stelmack, & Campbell, 1992) and to neural conduction velocity in the brain and periphery (Vernon & Mori, 1992). Also, processing rate measured in simple perceptual-motor tasks has been found to be related to capabilities on more complex cognitive tasks, in children

(Deary, 1995; Fry & Hale, 1996; Hale & Jansen, 1994) and across the life span (Kail & Salthouse, 1994; Salthouse, 1996). The data have been explained on the basis of a global rate parameter that differs among individuals and changes with development.

It is quite possible that at least some such investigators would agree that there could be more domain-specific processing rates in addition to a global rate, but the emphasis in the recent literature has been on a single, global rate. Similarly, it is quite possible that at least some such investigators would agree that there could be discrepancies between the causes of age effects and of individual differences (a point raised by Engle, 1996), but the emphasis in recent literature has been on the relation of both to a global processing rate.

An interesting test case for the application of the global processing rate principle is verbal short-term memory span. Span is an interesting task because of its simplicity: It measures how many items a person can repeat back in sequence. It is interesting also because of its widespread use in intelligence tests and its growing use in practical settings, such as personnel selection (Verive & McDaniel, 1996). It is one of the simplest tasks in which participants must hold information temporarily in a working memory in order to respond correctly (Baddeley, 1986) and should reveal basic characteristics of working memory that are applicable to more complex mental tasks. Phonologically based aspects of memory span in children have been shown to be relevant to vocabulary acquisition (Gathercole & Baddeley, 1993) as well as other language and reading abilities (Gillam, Cowan, & Day, 1995; Sipe & Engle, 1986).

There is ample evidence of a linear relation between verbal memory span and the fastest rate at which a person can pronounce very short lists drawn from the set of items used to measure span. The correlation occurs no matter whether the variation in span is caused by individual differences, age differences, or item-length differences (Baddeley, Thomson, & Buchanan, 1975; Hulme, Thomson, Muir, & Lawrence, 1984; Naveh-Benjamin & Ayres, 1986;

Schweickert & Boruff, 1986). A well-known explanation (Baddeley, 1986, 1992) is that speeded articulation estimates the rate of covert verbal rehearsal, which in turn determines how many phonological representations can be kept active in memory concurrently in a repeating "phonological loop."

According to the global rate hypothesis, one would expect that rehearsal rate is just one manifestation of the global rate. Research has suggested that there is a relation between the global rate of processing and memory span and that most or all of this relation can be attributed in turn to the strong relation between the global rate of processing and the rate of articulation, both in child development (Kail & Park, 1994) and in adult aging (Salthouse & Coon, 1993).

This does not mean, however, that a global rate factor (or a more specific, rehearsal rate factor) alone provides an adequate account of memory span. Typically, much of the age-related and individual sources of variance in memory span are left unaccounted for by measures of rehearsal rate or processing rate. The use of the term *global* to describe the processing rate variable seems to imply uniformity across tasks. The notion of a global processing rate is akin to the general or *g* factor in studies of human intelligence, which accounts for correlations in individuals' performance across mental domains (see, for example, Gustafsson, 1984; Spearman, 1927). Yet, it seems plausible that there could be important individual differences in the rates of various mental processes that cannot be reduced to a single global rate, just as the studies demonstrating the *g* factor in intelligence also have obtained evidence of domain-specific factors. In fact, Kail and Park (1994) found that some of the effects of age on memory span cannot be accounted for by global or articulatory processing rate. They tentatively attributed the residual variability in span to qualitative factors such as knowledge and mnemonic strategies and therefore did not directly challenge the assumption that all processing rate variables reflect a single, underlying global rate. However, their theoretical framework might allow that there could be more specific processing rates in addition to a general or global rate.

Upon careful consideration of the basic theoretical model that has most often been applied to an understanding of memory span and other tasks thought to involve working memory (e.g., Baddeley, 1986, 1992), it seems to us that one would not necessarily expect a global rate factor alone to be adequate. The model includes relatively automatic mechanisms that are used to hold information temporarily (in the case of verbal material, the phonological loop), and also relatively effort-consuming, "central executive" processes that control the deployment of the automatic mechanisms. Even though memory span is a relatively simple task, both of these mechanisms would have to be involved. For example, the individual might use the phonological loop to retain some of the items and might use the central executive to hold several more items in a nonarticulatory form (Zhang & Simon, 1985), perhaps in the focus of attention (Cowan, 1988, 1993, 1994, 1995). Theoretically, it remains quite possible that variations in these two mechanisms could occur separately. For example, one individual might have an excellent phonological loop but mediocre central executive

processes, whereas another individual might show the converse pattern of abilities. Some research on individual differences in memory-span task performance could, in fact, be interpreted as demonstrating separable variation in the central executive and phonological loop processes (Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996).

In two experiments, our primary focus was on the correlations between tasks and the model of performance that could be built. In each experiment, two types of timing measure were included. One type of measure focused on rapid overt or covert articulatory rates and may serve as an index of the efficiency of phonological loop-based rehearsal processes. According to theory (Baddeley, 1986), a more rapid articulatory rate allows the transient phonological representation of the items to be refreshed more quickly, which allows more items to be kept active at the same time.

Another type of measure focused on the timing of short-term memory retrieval in two different types of task. Its theoretical role is uncertain, but it may serve as an index of the efficiency of central executive processes as they are applied to memory-span tasks. Teasdale et al. (1995, p. 554) stated that "To date, the central executive's most clearly specified function has been the control and coordination of performance." In short-term memory retrieval, central executive processes may be important for keeping the mental representations of to-be-recalled items active while keeping them separate from any intrusive, interfering information and while keeping track of which item is the appropriate one to be retrieved at a particular moment.

We anticipated that both types of measure would correlate with memory span. If the global processing rate account of span is sufficient, then there also should be fairly high correlations between the two types of timing measure. However, if the two types of measure are uncorrelated even though both correlate with span, this would suggest that the global rate account is insufficient and that the efficiency or rates of multiple, independent faculties must be taken into consideration. Experiment 1 was a developmental investigation with elementary school children of three ages, whereas Experiment 2 used adult participants and novel measures of retrieval and rehearsal. The procedures differed substantially but converged on a single verdict: that the global rate hypothesis is insufficient. A general processing rate may exist, but it must be supplemented by one or more domain-specific processing rates to account for memory span.

Experiment 1

Rationale Behind the Rapid Articulation Measures

In order to examine the global rate hypothesis, this developmental investigation refined the study of memory span in several ways. First, it included a more careful examination of the rapid-speaking task that has been found to correlate with span. In most prior developmental studies, that task has involved the repeated pronunciation of a short set of items several times in a row, without stopping, in order to created a response long enough to measure reliably with a stopwatch (e.g., Baddeley et al., 1975; Cowan et al., 1994;

Gathercole, Adams, & Hitch, 1994). However, with that method it is unclear if the developmental differences should be attributed to the time taken to pronounce each word in the set or the time or effort needed to plan and initiate each cycle of the repetition.

Hulme et al. (1984) examined a related issue. They measured the speech rate for single words and word triads, asking children to repeat each word or word triad 10 times in a row as rapidly as possible. Children reproduced word triads more slowly than words, but the relation of both measures to span was about the same. Hulme et al. also spot-checked the separate durations of words and silent pauses between the words in a triad. They found that it was the duration of words, rather than pauses between words, that changed with age in this speeded-pronunciation task.

Our measures of rapid speech articulation were designed in the same spirit, but we simplified the tasks further. In one task, the child was to count from 1 to 10 only once, as rapidly as possible, after receiving a response signal. In another task, a short list was presented on each trial followed by a response signal, after which the participant was to repeat the list only once. These tasks produced few interword pauses. However, a computer program capable of measuring durations on an oscillographic display was used to obtain accurate measurements of the duration of the interval from the response signal to the beginning of the child's repetition (the preparatory interval) and from the onset of the first word in the response to the end of the last word (the speaking duration). These measurements thus separated the effects of preparation time from the effects of speaking time. The rapid list-articulation task was similar to one used with adult participants by Sternberg, Monsell, Knoll, and Wright (1978), though they did not focus on individual differences or comparisons with span task performance.

A potential problem with rapid articulation tasks is that even the short sets of items used in such tasks might impose a memory load, which might slow down the repetition. This memory load might be more important for the younger children because a given list length is a higher proportion of span for these children than for older children. No such effect was obtained in the aforementioned study by Hulme et al. (1984), but they compared the performance of 8- and 10-year-olds, whereas we used a wider range of ages. To examine this possibility of a memory load effect on rapid articulation, each list was used for several trials. Specifically, each random list of digits was used for six trials in a row. If short-term memory is a factor in the rapid articulation task, then the age difference in repetition rate might be expected to diminish across the repetitions of a random list as it became established in long-term memory, reducing children's reliance on short-term memory.

Rationale Behind the Retrieval Measures

In addition to examining rapid articulation in more detail than most previous studies, we also examined a timing measure in the memory-span task itself, involving the duration of the spoken responses. For reasons to be explained, pauses within these spoken responses are to serve as estimates of the short-term memory-retrieval time. In the span task, responses are made at the child's own preferred, unspeeded pace. Cowan (1992) measured the duration of words and interword pauses in 4-year-old children's span-task responses, using only correct responses to subspan lists. It was found that, for lists of a particular fixed length, children who produced shorter interword pauses in correct repetitions also went on to achieve higher scores on the span task. Additionally, Cowan et al. (1994) found that older (8-year-old) children produced shorter preparatory intervals and interword pauses in a span task and, of course, achieved higher spans than 4-year-old children did. The differences were much smaller for the duration of words: Children who could speak more quickly did not necessarily produce individual words more quickly in the span task.

It is not clear exactly what processes take place during the interword pauses. Regardless of the mechanism, the interword pauses are interesting in that they are related to memory span but do not seem to arise from articulation-related factors. Cowan et al. (1994) used mono-, di-, and trisyllabic words and found no effect of the word length on the duration of interword pauses in the responses (whereas, understandably, there was a large effect of word length on the duration of words in the responses). This absence of a word-length effect within the interword pauses is striking given the ubiquity of word-length effects when rehearsal is being carried out (e.g., Baddeley et al., 1975; Schweickert & Boruff, 1986), suggesting that verbal rehearsal is not important during the interword pauses. However, if the durations of pauses amount to just one more indication of a common processing rate, as global processing rate theorists would suggest, then they should nevertheless be correlated with the articulatory rate factors that predict memory span.

It is helpful to consider additional clues to the processes that occur during interword pauses in the verbal span responses. Cowan (1992) and Cowan et al. (1994) both found that the pause durations covaried with list length. This suggests that children were in some way processing not just the next item, but the entire list (or, at least, a proportion of it) during each interword pause. One possibility is that some type of memory search that is not phonologically based is being carried out during the pauses. The child may be searching through the short-term memory representation of the list to determine which word to say next (cf. Sternberg et al., 1978). Consistent with this conjecture, studies of short-term memory search using Sternberg's (1966) probe procedure are analogous to the findings of Cowan et al. (1994) for interword pauses in span-task responses, in that they have failed to find an effect of word length on search rate (Chase, 1977; Clifton & Tash, 1973). Thus, the pattern of interword pause times that we have obtained could reflect memory-search processes that are not based on rehearsal, and these search processes could vary among individuals and become more efficient with development.

Relation of the Measures to the Articulatory Loop Model

It is possible to use past research to begin to relate measures to three hypothetical processes in memory span

discussed by Baddeley (1986): the phonological storage buffer, the articulatory process, and central executive processes. Clarifying the distinction between phonological storage and articulation, Schweickert, Guentert, and Hersberger (1990) and Hulme and Tordoff (1989) differentiated the mechanisms of the phonological similarity and word-length effects. Phonological similarity between items in a list impedes recall, but does not affect the rate at which items are pronounced. It presumably operates by making items less distinct from one another within the phonological store. Recall also is impeded when the words to be recalled are relatively long, but in this case longer words do take longer to pronounce. They presumably impede recall by slowing the articulation process. When one looks at phonological similarity and word-length effects in relation to the timing of recall in the span task, the distinction between them is reinforced. Cowan (1992) showed that the timing of responses in a span task is comparable for phonologically dissimilar and similar items, even though the span is shorter for the similar lists. In contrast, Cowan et al. (1994) showed that the durations of words in the response are affected by the length of the stimulus words, as one would expect.

The recent findings from our laboratory may reinforce the further distinction between the articulatory loop, on one hand, and central executive processes, on the other hand. In particular, in contrast to the absence of an effect of *word length* on the durations of silent periods between words in the memory-span response (interword pauses), these pause durations are related to *individual differences* in span (Cowan, 1992) and to *age* in childhood (Cowan et al., 1994). They therefore may reflect central executive processes but not articulatory processes. The longer silent periods observed in less advanced individuals may reflect slower or less efficient retrieval processes, presumably because of a less mature central executive.

In the present study, the duration of articulations in a speeded task and the duration of pauses in correct responses within the span task both have been compared to span. They are thought to reflect rehearsal and retrieval processes, respectively. Theoretically, if they both correlated with span and with each other, the global processing theory would receive additional support. However, if they correlated with span but not with each other, this would argue against the sufficiency of a global processing-rate notion. There still could be a general rate factor, but other, more specific rate factors would have to be posited as well. Other aspects of span and articulation also have been examined to gain a better understanding of these processes.

Method

Participants

The participants were 72 children (39 girls, 33 boys), with 24 children each in the first, third, and fifth grades. The ages of children in the three groups (in months) were 90.17 (*SD* = 5.71), 112.12 (*SD* = 5.32), and 136.40 (*SD* = 5.63), respectively. The selection of ages was limited at the low end by children's ability to last through the entire test session. At the high end, children's span begins to approximate the span of adults. It seemed desirable to

have all participants be in the elementary school years to maximize their shared environment.

Apparatus and Stimuli

Each child was tested individually in a sound-attenuated room for about 1 hr. The stimuli for all measurements were single-digit numerals spoken in a female voice, each lasting about 500 ms, randomly arranged into lists for most of the tasks. These stimuli were digitally stored and reproduced using a Macintosh II computer (Apple Computer, Inc., Cupertino, CA) and audiological headphones. Following each list was a 104-ms, 1000-Hz tone that served as a response signal.

Procedure

Span Test 1. A memory-span task was administered first. Each list was presented at a one digit per second rate, with the response signal following 1 s after the onset of the last digit in the list. Testing began with four random two-digit lists. On each trial, the child was to repeat the digits in order, at whatever pace he or she wished, immediately after the response signal. Then the list length was incremented by one and testing continued, until a length was reached at which three of four lists were repeated incorrectly. The digit 7 was omitted from the span and rapid list-articulation tasks in order to restrict the stimuli to monosyllabic words.

Speeded articulation tasks. Next, speeded articulation times were measured in two ways. First, there were two trials for the child to count from 1 to 10 as rapidly as possible. Second, in the rapid list-articulation task, a spoken list of digits (presented at a 1 digit per second rate) was followed by a tone 1–1.5 s afterward, signaling the child to repeat the list once as soon and as quickly as possible. Each child received a different random list with 1, 2, 3, and 4 digits. There were six successive trials with each of the four lists. The order of list lengths in this task was counterbalanced across children. This task was followed by another two trials of counting rapidly from 1 to 10.

Span Test 2. Finally, there was another run through the memory-span task with different digit randomizations.

Measures and Types of Analyses

Memory span. Memory span for each child was defined as the length at which all four trials were answered correctly plus a 0.25-item credit for each correct trial at a higher length.

Timing measurements. Timing measurements for the span-task responses and for responses in the rapid articulation tasks were made on the basis of tape recordings from the test session, using an oscillographic display on the computer to determine each response segment's duration to the nearest 10 ms. Only correctly repeated lists were used. In the span-task responses, measures of the duration of interword pauses were considered most important, as an index of the time taken for memory retrieval (Cowan, 1992; Cowan et al., 1994), but measures of preparatory intervals and word durations in the response also were made, in order to yield clues about processing. The measurements were averaged across lists of a particular length for each child. Measurements were taken for List Lengths 2, 3, and 4, with higher list lengths omitted because not all children correctly repeated any lists of Length 5 or higher. On the basis of a visual inspection of a histogram of the responses, which showed a unimodal distribution of responses up to about 1.8 s, trials with any one silent pause longer than 1.8 s were considered outliers and were omitted. This occurred in only eight trials. In the rapid articulation tasks, the preparatory intervals and the total

speaking duration were measured separately for each trial. Unlike the span task, speaking duration in the rapid articulation tasks was not subdivided into words and pauses because there were few measurable pauses.

ANOVAs. In all analyses of variance (ANOVAs), all factors except age group were within-subject factors. Separate ANOVAs were carried out for each of the dependent measures (memory span, each of three measures of response timing in the span task, and rapid list articulation).

Correlations. Intercorrelations among all of the main measures were calculated.

Latent variables and structural equation modeling. Because we had multiple empirical measures of three basic theoretical mechanisms (short-term memory-retrieval rate, represented by the durations of interword pauses in span-task responses; rehearsal rate, represented by the durations of rapid articulations in the list and counting tasks; and short-term memory span, represented by two span runs), we were able to use a latent variable, structural equation model (Hoyle, 1995) to determine how the retrieval and rehearsal mechanisms might account for variance in short-term memory. This type of analysis uses the common variance between multiple measures of a particular underlying theoretical construct, or *latent variable,* to estimate the quantitative contribution of that construct. The latent variables are combined within a structural equation model specifying the hypothetical relation between these variables, and the strength of these hypothetical relations can be assessed. Unlike simple regression equations, a latent variable model attempts to account not for all variance, but only for the variance that is shared between multiple measures of a particular theoretical construct and therefore, presumably, is theoretically meaningful. Latent variable analysis was conducted with the CALIS procedure within the SAS statistical package (SAS Institute, Inc., 1989). Preparatory intervals in all tasks and word durations in the span-task responses were excluded from these analyses because their relation to span is unclear theoretically (Cowan, 1992; Cowan et al., 1994).

Results and Discussion

The primary hypothesis under examination, the global processing rate hypothesis, is addressed primarily by the correlations and latent variable model. However, means and ANOVAs are reported first in order to examine processing in the memory-span and rapid-articulation tasks. The main contribution of these analyses is to show that the responses in this experiment (a) are generally consistent with previous studies, (b) change systematically with age, and (c) serve as evidence on the nature of several covert processes in short-term recall.

Means and ANOVAs

Memory span. Table 1 presents the main measures for all tasks as a function of age group (and list length where applicable), along with an indication of which group main effects were significant. As shown in this table, memory span increased systematically across ages. There was no significant difference between Span Tests 1 and 2.

Interword pauses in the memory-span task. We were able to obtain correctly repeated lists in all children for List Lengths 2, 3, and 4, so our response timing analyses are based on these list lengths. The main timing measure from the span-task response was the duration of interword pauses, which Cowan (1992) suggested was an index of retrieval processes that contribute to span-task performance. Table 1 shows that in this experiment, consistent with this expectation, the duration of these interword pauses for a particular list length decreased with age, presumably because the rate or efficiency of retrieval increased.

The duration of interword pauses increased across list

Table 1

Mean and Standard Deviation of Each Measure by Age Group in Experiment 1

| | Grade in school | | | | | | | |
| | First | | Third | | Fifth | | Age effect | |
Measure	M	SD	M	SD	M	SD	F(2, 69)	MSE
Memory span estimates							10.58*	0.65
Span Test 1	4.53	0.70	4.85	0.66	5.17	0.62	—	—
Span Test 2	4.49	0.66	4.87	0.56	5.37	0.55	—	—
Interword pauses in the								
memory span task (ms)							7.65*	41,521
List Length 2	173	92	149	117	99	102	3.08	10,953
List Length 3	257	132	202	135	121	148	5.86*	19,125
List Length 4	341	172	264	141	148	160	9.09*	25,063
Speaking durations in the rapid								
list articulation task (ms)[a]							3.75*	40,567
List Length 1	225	78	190	69	190	67	1.92	5,144
List Length 2	354	90	345	84	346	82	0.08	7,262
List Length 3	544	166	499	98	493	84	1.24	14,784
List Length 4	845	337	708	113	632	120	5.99*	46,891
Speaking durations in the								
rapid counting task (ms)[b]	1,705	212	1,497	228	1,475	193	8.61*	178,882

Note. Statistical tests listed on the same line as a major heading included the data for all of the subsidiary headings shown (i.e., for both span tests or for all list lengths). Dashes indicate that tests were not conducted.
[a]Collapsed across Repetitions 1–6 for each list length. [b]Collapsed across Repetitions 1–4.
*$p < .05$.

lengths, $F(2, 138) = 32.24$, $MSE = 6,810$, $p < .001$. This list-length effect might be understood on the grounds that the child must mentally search through the list representation in short-term memory to determine which digit to pronounce next. This search process may take longer when the list is longer (cf. Cowan, 1992; Sternberg et al., 1978). Additionally, there was an interaction of Age × List Length, $F(4, 138) = 3.20$, $MSE = 6,810$, $p < .02$. As Table 1 shows, the list-length effect was more severe in the younger children. This interaction would be expected if the search process during each pause entails a search through the entire list or a proportion of it. In that case, the younger children's search disadvantage would be compounded as the list length increased. Alternatively, given that lists in the span task were presented in an ascending order (the usual procedure), there could have been more fatigue for longer lists, especially in younger children.

Speaking durations in the rapid list-articulation task. Responses for this rapid articulation task were rather orderly and comparable in form to what Sternberg et al. (1978) previously had obtained (using a similar task, but with adult participants, longer list lengths, and more practice). The speaking durations rather than the preparatory intervals were of primary interest to us given that these comprise a relatively precise version of the type of measure that has been used to estimate rehearsal rate ever since the work of Baddeley et al. (1975) and given also the evidence from Hulme et al. (1984) that word durations, rather than interword intervals, within speeded articulations are responsible for their correlation with span. In our data, the articulatory response durations in rapid list articulation were orderly, as shown in Table 1. There was a significant effect of list length, $F(3, 207) = 329.35$, $MSE = 67,028$, $p < .001$. The mean speaking durations for List Lengths 1–4 (across all list items) were 202, 348, 512, and 729 ms, respectively. There also was a significant effect of repetition, $F(5, 345) = 61.88$, $MSE = 10,678$, $p < .001$. The mean speaking durations on Repetitions 1–6 for a particular list length were 539, 456, 439, 422, 417, and 412 ms, respectively. This constitutes a practice effect. The form of the durations across list lengths is quite similar to what was observed in a comparable study with adults (Sternberg et al., 1978), though the present durations are considerably longer.

Additionally, there was a List Length × Repetition interaction, $F(15, 1035) = 13.71$, $MSE = 10,512$, $p < .001$. One simple way to describe this interaction is by calculating a slope of speaking times across Repetitions 1–6, separately for each list length. The slopes for List Lengths 1–4 were -2.14, -10.91, -27.69, and -47.46 ms per additional repetition, respectively. This indicates that practice effects were larger for longer lists, as one might expect.

Table 1 shows that there was an overall effect of age on the speaking durations in this task. However, as the table suggests, there also was a consistent Age Group × List Length interaction, $F(6, 207) = 4.70$, $MSE = 67,027$, $p < .001$. As shown in Table 1, the age group effect was significant only for four-word lists. Apparently, speaking durations are more heavily influenced by the memory load imposed by longer lists in the younger children. Unlike an

analogous effect for interword pauses in the span task, the present Age Group × List Length interaction cannot be attributed to fatigue, inasmuch as list lengths were randomized in this task. Finally, there was an interaction of Age Group × Repetition, $F(10, 345) = 6.14$, $MSE = 10,678$, $p < .001$. One simple way to describe this effect is to calculate a slope for performance across repetitions. This slope was -74.57, -42.71, and -24.97 ms per repetition for the three age groups. Thus, younger children benefited from practice more than older ones did.

Speaking durations in the rapid counting task. In the analysis of the speaking durations in the counting task, there was an age effect (see Table 1). There also was an effect of repetition, $F(3, 207) = 21.87$, $MSE = 15,480$, $p < .001$. For the four repetitions the means were 1,633 ms, 1,600 ms, 1,489 ms, and 1,514 ms, respectively. In pairwise comparisons, the first two repetitions took longer than the last two, $ps < .01$. Thus, there was improvement with practice, which was approximately equivalent across age groups. Recall that an age effect also was obtained for random lists of four digits, though not for shorter lists.

Auxiliary timing measures. Additional, auxiliary timing measures can help to complete the description of performance in all tasks. The results for such measures are described and discussed in Appendix A. These measures include preparatory intervals in all tasks and the duration of words in the span-task responses. Serial-position effects in span-task responses also are considered within this appendix. The results show that these measures, like the main measures, changed in an orderly way with age and offer clues to the processes involved in memory-span performance. However, Appendix A provides further explanation for why these measures were not considered appropriate for the latent variable model analysis of memory span. The reasons stem both from prior research and from our findings reported in the appendix.

Summary of means and ANOVAs. There was an orderly increase in span and decrease in various latencies and response durations with age. The detailed pattern of results suggests that both articulation and various covert mnemonic processes generally become faster and more efficient as children mature.

Correlations Among Measures

The following analyses allow a better understanding of the sources of variation in the memory-span task and therefore are relevant to the issue of whether only a global processing rate or multiple, independent rates are involved. They also are relevant to the question of whether the influences of age and individual differences on span are similar or different. A further question, to be addressed with a latent variable model, is how the different variables that have been examined may operate together in the memory-span task.

Main measures. Table 2 presents the correlations among all of the main measures (age, span measures, and the timing variables that we expected to be related to span on the basis of prior research). First, note that memory span and age were

Table 2
Correlations Among the Main Measures in Experiment 1

Measure	Age	MSp	Sp1	Sp2	Ps2	Ps3	Ps4	Ra1	Ra2	Ra3	Ra4
Mean Span (MSp)	.43**	—									
Span Test 1 (Sp1)	.32**	.93**	—								
Span Test 2 (Sp2)	.47**	.92**	.71**	—							
Interword pauses in the memory span task											
Length 2 (Ps2)	−.29**	−.26	−.25	−.23	—						
Length 3 (Ps3)	−.34**	−.40**	−.33**	−.41**	.78**	—					
Length 4 (Ps4)	−.41**	−.47**	−.41**	−.46**	.60**	.80**	—				
Speaking durations in the rapid list articulation task											
Length 1 (Ra1)	−.21	−.21	−.18	−.20	−.03	.12	.15	—			
Length 2 (Ra2)	−.03	−.39**	−.44**	−.28**	−.09	−.04	.01	.54**	—		
Length 3 (Ra3)	−.15	−.37**	−.36**	−.33**	.05	.07	.06	.40**	.57**	—	
Length 4 (Ra4)	−.32**	−.53**	−.49**	−.49**	.17	.24	.18	.35**	.50**	.67**	—
Speaking durations in the rapid counting task	−.39**	−.33**	−.27	−.34**	.00	.04	.16	.31**	.43**	.49**	.49**

Note. Notice that the interword pause measures were intercorrelated, as were the rapid articulation duration measures, but that these two groups of timing measures were unrelated to one another.
**$p < .01$, one-tailed.

correlated at a moderate, though significant, level, $r = .43$. This implies that over .81 ($= 1 - .43^2$) of the variation in span reflects not age effects, but instead either stable individual differences at each age or intraindividual instability in the span test. Notice also, though, that Span Tests 1 and 2 were fairly highly correlated, $r = .71$. This correlation suggests that much of the non-age-related variability in span resulted from individual differences within each age, as the span tests were fairly reliable.

Second, it can be observed in Table 2 that the timing measures generally are correlated with span measures according to the criterion we adopted ($p < .01$, one-tailed). Interword pauses in the span task were correlated with span, though the correlations were not quite significant for the shortest (two-word) lists. Similarly, the duration of speech in the rapid articulation task was correlated with span, though the correlations did not reach significance for the shortest (one-word) lists.

The table shows that the interword pauses for various list lengths were significantly intercorrelative ($.60 < r < .80$), as were the rapid articulation responses for various list lengths ($.35 < r < .67$). Yet, despite this internal consistency of each timing measure, and despite the correlations between both types of measure and span, the table also shows near-zero correlations between interword pauses and speaking durations in the rapid articulation task ($-.09 < r < .24$, mean r across nine correlations $= .07$). This indicates that the two types of timing measure must be picking up different portions of the variability in span, raising the hope that the two types of timing variable together could account for much of the span variance.

Auxiliary measures. Appendix B presents correlations involving auxiliary timing variables (word durations in the span task and preparatory intervals in all tasks) that were not expected to be related to span on the basis of previous work (Cowan, 1992; Cowan et al., 1994) and were not included in

the latent variable model. Prior expectations were not entirely met, inasmuch as the duration of words within correct list repetitions in the span task were, in fact, generally correlated with span. However, Appendix B shows that word durations also were correlated with interword pauses ($.53 < r < .70$). Perhaps the processing taking place during word pronunciations sometimes includes some of the same memory-search processes that are proposed to take place during the interword pauses. This concurrent processing during word productions in the span task could slow down the pronunciation of some words.

We also looked at a composite measure in span responses, the time from the beginning of the first word in the response to the end of the last word. Its correlations with span for List Lengths 2, 3, and 4 ($rs = -.24, -.32$, and $-.44$, respectively) were quite similar to the correlations between span and word durations (Appendix B), and somewhat smaller than the correlations between span and interword pauses (Table 2: $rs = -.26, -.40$, and $-.47$). Inasmuch as the composite measure includes all of the word durations and interword pauses in the response, it seems clear that the word durations do not make an important, unique contribution to the understanding of span beyond what one obtains from the interword pauses. Similarly, stepwise regressions predicting span from interword pauses showed $R^2 = .230$, but there was an insignificant increase in R^2 (to .235) when the durations were added to the equation. These findings and the theoretical ambiguity of word durations were behind our decision not to include word durations in the latent variable model.

The correlations involving preparatory intervals in the rapid articulation task, shown in Appendix B, proved to be relevant to the question of the relation between age effects and individual differences. The correlations between span and preparatory intervals in this task were not significant ($rs = -.11, -.11, -.21$, and $-.05$ for List Lengths 1–4,

respectively). In contrast, the correlations between age and preparatory intervals were all significant ($rs = -.38, -.34, -.39,$ and $-.35$). Conversely, one might point to the rapid articulation durations themselves, which, at List Lengths 2 and 3, correlated with span ($-.39$ and $-.37$) but not with age ($-.03$ and $-.15$), as shown in Table 2. Such dissociations between age and span are possible because age and span share only a minor portion of their variance, $R^2 = .18$. Engle (1996) similarly noted the importance of distinguishing between sources of variance due to age versus individual differences within each age, the point being that they are not necessarily influenced in the same way by particular processing variables.

Summary of correlations. The outcome of the correlational analyses was basically as predicted, though there was some additional new information among the correlations. The correlations between the span-task interword pause durations and memory span in that task (see Table 2) were predicted on the basis of previous studies of this measure (Cowan, 1992; Cowan et al., 1994). The correlations between speaking durations in the speeded tasks and memory span would have been predicted from a wealth of previous literature (e.g., Baddeley et al., 1975; Hulme et al., 1984; Naveh-Benjamin & Ayres, 1986; Schweickert & Boruff, 1986) as discussed on pages 141–142. The relative weakness and nonsignificance of these correlations for the shortest list length in each case was not expected, but it did not contradict previous research, either. Apparently, a certain minimal memory load is needed before these tasks are relevant to span.

What was especially interesting about these correlations with span is that the 12 correlations between interword pauses in the span-task responses, on one hand, and durations of productions in the rapid articulation tasks, on the other hand, all were nonsignificant despite consistently significant correlations within each class of timing variable (see Table 2). It appears that the two types of timing measure pick up different portions of the variance in span.

There were no clear predictions for the auxiliary timing measures summarized in Appendix B. The preparatory intervals proved to be uncorrelated with span in any of the tasks despite some age effects. (The converse pattern occurred in the rapid articulation task among the speaking durations for List Lengths 2 and 3, which showed no age effects but were correlated with span nevertheless.) The duration of words in the spoken responses within the span task did unexpectedly show sizable correlations with span. However, these word durations appeared to account for little or no variance in memory span that was not also reflected in the interword pauses, given that neither word durations nor a composite measure across words and pauses (speaking duration) correlated with span as strongly as did the pauses alone.

Latent Variable, Structural Equation Modeling

Our basis for a structural equation modeling of performance was a small set of latent variables formed from the measures that were deemed relevant on theoretical grounds,

as discussed previously. A rehearsal rate variable was based on the mean speaking times in the rapid articulation tasks, including the duration of counting and of responding to random number lists. A retrieval rate variable was based on the mean duration of interword pauses within correctly repeated lists. The memory-span latent variable was composed of the two span runs. In the model, age could influence the durations of these two processes, and these durations in turn could influence memory span. Finally, age could influence memory span directly. Our final model of this process is shown in Figure 1.

The model shown in the figure is based on separate timing measurements at each list length. However, in the span task, the interword pauses for the shortest list length (Length 2) were not reliably correlated with span (see Table 2) and were omitted from the model. Similarly, in the rapid list-articulation task, the speaking durations for the shortest list length (Length 1) were not reliably correlated with span and were omitted from the model. Although the low correlations for these shortest list lengths were not anticipated, it makes sense that a certain minimal memory load might be needed before the timing measures are highly relevant to span. (The undesirable consequences of including these measures in the model are discussed a little later.)

In Figure 1, notice that both of the path coefficients from age to the rate variables are significant, as are both of the path coefficients from the rate variables to span. The coefficient for the path directly from age to span is not significant. In each case, the coefficient represents "the change in the dependent variable, in standard deviation units, expected to result from a change of one standard deviation in the causal variable, if other possible determinants specified in the model were held constant" (Fry & Hale, 1996, p. 239). The path coefficients involving rate variables are negative because the rate measures were expressed in terms of duration per item rather than its reciprocal. Other variables in the formal model that are not shown include error terms for each measured variable except age and, analogously, disturbances for each latent variable. The correlation between the disturbances for the two rate-latent variables was included in the model but was negligible (.05), reinforcing the assumption that the processes underlying these latent variables operate independently of one another.

The goodness of fit to the data of the model shown in Figure 1 was examined with various alternative, conventionally preferred statistics (recommended, for example, by Marsh, Balla, & McDonald, 1988). The chi-square fit was $\chi^2(22) = 35.74$. This significant chi-square statistic indicates that the model still does not fit the data perfectly, but that is the usual state of affairs in modeling approaches: The chi-square fit is considered overly sensitive to small discrepancies (Marsh et al., 1988) but is usually reported anyway.

Certain other fit indices are generally considered more diagnostic, though there is no general agreement as to which index is best. To mention a few of the most often-used fit indices for our model, the goodness-of-fit index was .90, the Bentler-Bonett Normed Fit Index was .88, and Bentler's Comparative Fit Index was .95. A goodness-of-fit index of .9
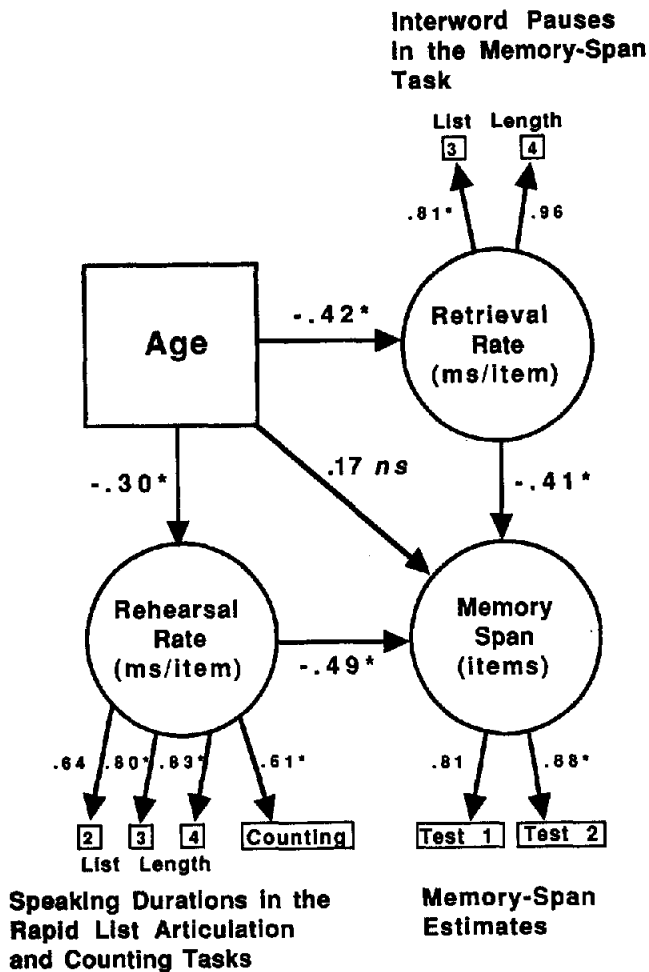
**Interword Pauses
In the Memory-Span
Task**

Figure 1. A latent variable model accounting for children's memory-span performance in Experiment 1. Error and disturbance paths have been omitted from the figure for the sake of simplicity. Error variances and one path from each latent variable to an empirically measured variable were set to 1.0 in order to obtain an identified solution. The covariation between disturbances for the two rate latent variables was near zero (.05), indicating their nearly total independence. The path coefficients (numbers shown along the arrows) indicate that an individual's span can be predicted unusually well from the retrieval and rapid articulation measures taken together. Coefficients for paths to and from the rate variables were negative because rates were expressed as durations per item rather than its reciprocal. Asterisks (*) denote significant path coefficients, $p < .05$.

or above on any of these measure is considered very good (Loehlin, 1992).

The model shown in Figure 1 accounted for 60% of the variance in span. (This proportion is calculated as 1.0 minus the square of the residual error of the memory-span latent variable.) However, when these paths from the timing measures to span were disallowed in a null model, the coefficient directly from age to span rose to a significant .47, and the percentage of variance in span accounted for by the model declined to only 22%. Thus, 38% of the variance in

span (= 60% - 22%) was related to the rate variables but not to age.

The model also indicated that the rate variables accounted for as much as 87% of the variance in the unadjusted age–span relation. That percentage was derived from the variabilities associated with the age–span path coefficients in the full model (.17) and the null model, with the paths from rate variables to span disallowed (.47): $(.47^2 - .17^2)/.47^2 = .87$. In short, the analyses indicated that there are two quite separate rate factors related to memory span and that the two of them together account for a large proportion of the variance in span. The significance of the path coefficients from both rehearsal rate and retrieval rate to span (shown in Figure 1) logically implies that both sets of contributing variables are needed and that the model would be significantly poorer if either set of variables was not allowed to contribute to span.

Another, control model was constructed to help assess the validity of the assumption that two independent rate variables contributed to span. In this control model, all of the timing measures contributing to the model shown in Figure 1 were still included, but they were now assumed to contribute to a single *processing-rate latent variable* rather than two distinct variables. This model accounted for 50% of the variance in span, 10% less than the 60% accounted for by the full model. The fit of this control model, $\chi^2(25) = 120.15$, was significantly poorer than that of the full model described previously, difference $\chi^2(3) = 84.41$. The path coefficient from the general processing-rate latent variable to span was −.73, and the one directly from age to span remained significant at .24, unlike the full model in which the age–span coefficient became nonsignificant. The link from age to the general rate variable was −.21. It is clear from these results that separate rehearsal and retrieval latent variables improve the fit in comparison to a single processing-rate latent variable, even when the same empirical measures are included in the model.

Finally, we also tried another model that was like the one shown in Figure 1, except that it included the two timing measures that were not significantly correlated with span as shown in Table 2 and had been omitted (Length 2 interword pauses in the span task and Length 1 rapid list-articulation times). The results of this second model were similar to the model shown in Figure 1. This model accounted for 54% of the variance in memory span. The path coefficient linking age and the rate of retrieval was −.35; linking age and rehearsal rate, −.29; and linking age directly to span, .24. The link between retrieval rate and span was −.32, and between rehearsal rate and span, −.46. The correlation between error variances (disturbances) for the retrieval and rehearsal latent variables was only .04. The latent variable model, however, provides additional diagnostic information about the suitability of the model. In this model, the estimate of error variance of one manifest variable (the first span run) was −.61. A negative variance estimate is nonsensical and is a sign that the model is underdetermined, so the estimates of the paths between latent variables in this model cannot be considered accurate. (Possibly, this problem would have been eliminated with a larger sample size, which was limited

by the tedious nature of the timing measurements.) Thus, the model parameters shown in Figure 1 are likely to be more accurate.

Experiment 2

The results of the developmental experiment suggested that two different types of rate measures together account not only for age effects in span but also for a large amount of the interindividual variance in span within each age. The two latent variables summarizing these timing measures were tentatively named *rehearsal rate* and *retrieval rate* on theoretical grounds. However, it could be argued that we used only indirect measures of these theoretical concepts. Conceivably, there could be factors that affect the rate of recall during the span task other than retrieval from short-term memory. It also is possible that the factors that contribute to performance in the rapid pronunciation task are not identical to those that occur during covert rehearsal, although Landauer (1962) has demonstrated a rough equivalence between the two.

In this second experiment, the theoretical approach was extended in several ways. We used college students rather than children as participants, in order to learn whether the model could apply to this group as well. Additionally, we devised more direct measures of the theoretical concepts. It clearly would be impressive if a similar model fit the data despite these drastic changes in method.

The methods were devised in a way that permitted group testing. In the span measure, participants were to listen to each list of numbers or letters and then write it down verbatim. The process continued up through list lengths that ensured that almost all participants would make errors, preventing ceiling effects.

In the measures of rehearsal rate, participants silently recited (rehearsed) well-learned sequences of items (the alphabet or the number series 1–10) repeatedly, making a mark on the response sheet to indicate each initiation of a rehearsal cycle. The period of rehearsal was held fixed, and the number of recorded cycles or repetitions in that period served as a measure of the rehearsal rate. In using this method one relies upon the participant for an accurate report, but if a correlation with memory span is nevertheless obtained, the usefulness of this type of rehearsal measure is clear.

The measures of retrieval from short-term memory were obtained not in the span task, as in the previous experiment, but rather in a separate task. It was modeled after the probe reaction time procedure of Sternberg (1966), though the processes involved are not entirely comparable to that well-known procedure. In each memory-search trial of our experiment, the participants listened to a set of one, three, or five items that were to be considered the target set. Then, for a fixed period of time, they viewed an array of items and circled as many of the targets as they could find within the array. The number of items circled in each condition was converted to a measure of rate in milliseconds per item, and the measure of memory-search efficiency was the slope of the function relating rate to memory set size. This slope

served as an indication of how much the rate of search changed with increases in the memory set size. The slope in this task was considerably higher than Sternberg (1966) and others typically have obtained, probably because in our modified procedure a recently presented memory set had to be maintained in the face of interference from nontarget items in the array. However, our design seems more similar to retrieval in a memory-span situation because, in that situation, items must be retrieved from the memory set in the face of output interference (Cowan et al., 1992). A similarly high slope was obtained by Sternberg (1969) in an experiment in which a separate short-term memory load had to be maintained during the memory-search task.

Method

Participants

The participants were 172 college students (99 women, 71 men, and 2 who did not note their gender) receiving credit for their introductory psychology course. They were tested in groups of about 20. Nine additional participants were excluded because of anomalous or incomplete answer sheets.

Stimuli and Procedure

The digit span test was first. The stimuli were read aloud by the experimenter, and each list was followed by a green card that served as a silent cue for a 10-s-long, written recall period to begin.

Each participant completed two trials with four-item lists and four trials at each of five different list lengths: 5, 6, 7, 8, and 9. List lengths were presented in a steadily increasing order, with randomly selected items in each list. (Unlike Experiment 1, the digit 7 was included). Next, the same procedure was carried out for a letter-span task, with the items drawn from the set [C, F, H, J, L, Q, R, T, Y].

The short-term memory-search tasks followed. In each task, the memory set was spoken twice by the experimenter. Then the participant had 30 s to circle as many of these memory targets as possible within a 15 × 16 array of targets and foils. This task was carried out six times: with memory sets composed of one letter (R), one digit (7), three letters (L, Y, C), three digits (8, 3, 5), five letters (Q, H, F, J, T), and five digits (2, 9, 1, 6, 4), in that order. The foils were other digits for the digit memory trials and other consonants for the letter memory trials, chosen so as to minimize the number of rhymes. Sixty of the 240 items in each array (25%) were target items. A different array was used for each memory-search trial.

The relation of this task to the more traditional search task of Sternberg (1966) was examined in a separate participant sample (N = 32), as reported in Appendix C. It appears to involve a relatively slow, self-terminating search (i.e., ending when the target is located), similar to what one might expect to occur when people conduct searches during a memory-span response but different from the rapid, exhaustive search (i.e., always running through the entire list) that Sternberg (1966) reported. Sternberg (1975) reviewed evidence that processes that slow down the search do tend to produce a self-terminating search rather than an exhaustive search.

In the first rapid rehearsal task, participants had 1 min to rehearse the alphabet silently, over and over as quickly as possible, making a pen mark on the page each time the alphabet began. In the second task, participants had 30 s to rehearse the numbers 1–10 silently, similarly marking the beginning of each counting cycle. In these

tasks and the remaining ones, the response periods began when the experimenter said "go" and ended when the experimenter said "stop."

Results and Discussion

Means and Correlations

The mean and standard deviation of each measure in the experiment is reported in Table 3. Table 4 shows the intercorrelations among all measures. The first point of note is that there was a relatively high correlation between the digit and letter memory-span measures ($r = .67$) and between the digit and letter rehearsal measures ($r = .69$). These correlations might be taken as conservative estimates of reliability, limited by any differences in how letters versus digits are processed. The correlation between the digit and letter memory-search measures was lower (.32), but still significant.

Both memory-search slopes were significantly correlated with both span measures, $-.20 > r > -.28$. Similarly, both rehearsal rate measures were significantly correlated with both span measures, $.23 < r < .25$. Because rates are expressed as seconds per item, unlike Experiment 1, the correlations are positive.

These correlations are consistent with previous work. The correlation between speech rate and memory span is consistent with Baddeley et al. (1975) and a large literature discussed on pages 141–142. Short-term memory-search rate tasks have been compared to memory span in a few studies, also. Cavanagh (1972) found a linear relationship between the two tasks when examined across the means for different materials. Puckett and Kausler (1984) tested nine different correlations in individual participants between span and the slope of memory-search rates in Sternberg's (1966) procedure. They obtained correlations that averaged .28 (the highest being .64). Most of these were not significant, though, given that only 20 individuals were tested. Our experiment establishes this correlation between short-term memory-search rate and memory span more soundly, though with a modified version of the memory-search task.

Table 3
Mean and Standard Deviation of Each Measure Used in Experiment 2

Task	M	SD
Memory span estimates		
Digit span	7.10	0.87
Letter span	6.33	0.89
Short-term memory search task slope		
Digit search task (slope in ms per memory set item)	190.52	84.50
Letter search task (slope in ms per memory set item)	264.04	130.97
Rapid covert repetition rates		
Rapid counting rate (silent repetitions of 1–10 in 30 s)	18.18	3.67
Rapid alphabet rate (silent repetitions of A–Z in 60 s)	13.73	3.00

Table 4
Correlations Between Measures in Experiment 2

Measure	DSp	LSp	DAr	LAr	Cnt
Memory span estimates					
Digit span (DSp)	—				
Letter span (LSp)	.67**	—			
Short-term memory search task slopes for arrays (Ar)					
Digit search (DAr)	−.27**	−.28**	—		
Letter search (LAr)	−.22**	−.20**	.32**	—	
Rapid covert repetition rates					
Rapid counting rate (Cnt)	.23**	.25**	−.08	−.01	—
Rapid alphabet rate (Alp)	.25**	.23**	−.08	.02	.69**

Note. Notice that the search slope measures were intercorrelated, as were the rapid rehearsal rate measures, but that these two groups of timing measures were unrelated to one another.
**$p < .01$, one-tailed.

An important finding is that the correlations between the rehearsal and retrieval (search) measures were nonsignificant, $-.08 < r < .02$. As in Experiment 1, the pattern is one in which two types of timing measure correlated with span, but not with each other.

Latent Variable, Structural Equation Model

A latent variable model for the key results is shown in Figure 2. The model was constructed and path coefficients were estimated by the same procedure as that used in Experiment 1. Once again, the model indicated that short-term memory-retrieval and rehearsal factors are independent of one another and that both are related to memory span. For the model shown in Figure 2, $\chi^2(1) = 0.40$, the goodness-of-fit index was 1.00, the Bentler-Bonett Normed Fit Index was 1.00, and Bentler's Comparative Fit Index was 1.00. These statistics indicate an excellent model fit. The model accounts for 40% of the variance in span.

In a control model, only one processing-rate latent variable was allowed, with all of the timing measures from the full model contributing to it. This measure accounted for only 18% of the variance. The coefficient from the general timing measure to span was .42, and the fit of the model, $\chi^2(3) = 32.39$, was significantly worse than before, difference $\chi^2(2) = 31.99$.

General Discussion

This research is relevant to several different issues that are addressed in turn. First, we have reviewed the results and
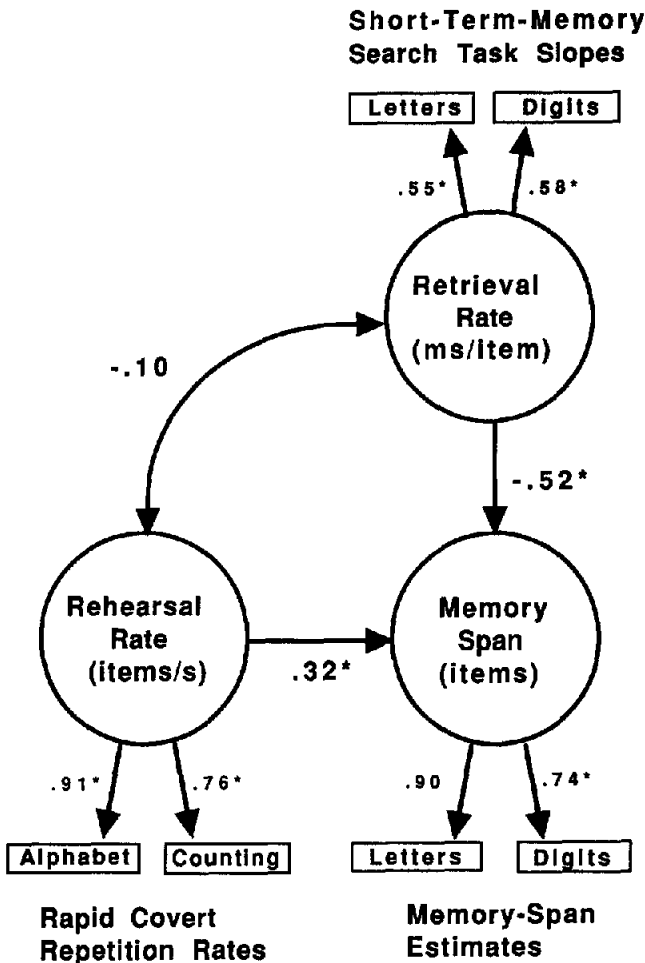
**Short-Term-Memory
Search Task Slopes**

Figure 2. A latent variable model accounting for adults' letter-span performance in Experiment 2. Error and disturbance paths have been omitted from the figure for the sake of simplicity. Variances of exogenous factors and the path from memory span to letters were set to 1.0 in order to obtain an identified solution. Coefficients for paths to and from the retrieval rate variable were negative because rates were expressed as durations per item rather than its reciprocal. Asterisks (*) denote significant path coefficients, $p <$ .05. The model accounts for 40% of the variance in span.

their implications for the standard articulatory loop model of memory span. Second, we have examined implications for more general cognitive models in which a global rate of processing is said to influence performance on many tasks, of which memory span is but one example. Third, we have reconsidered the relation between age effects and individual differences, which often have been assumed to stem from the same sources of variance. Fourth, on the basis of the results we have suggested a modified model of span. Fifth, and finally, we have discussed remaining questions and uncertainties.

Reevaluating the Articulatory Loop

The results clearly show that the leading model of working memory, the phonological loop model (Baddeley,

1986, 1992; Gathercole & Baddeley, 1993), has merit, but is an oversimplification. In addition to the rate of rehearsal that is critical within the model, our results suggest that the rate at which items are retrieved from short-term memory also plays an important role. In Experiment 1, the duration of articulation in a speeded task (presumably a measure of rehearsal processes) was shown to be uncorrelated with the duration of pauses within the responses in correct list repetitions in the memory-span task (presumably a measure of retrieval processes). Yet, both of these types of measure were correlated with memory span. In a latent model of processing, the two types of measure made different, unique contributions to span, and the model accounted for 60% of the overall variance in span and 87% of the age-related variance in span. In Experiment 2, a similar modeling result was achieved in a larger adult sample, using covert recitation as the rehearsal process measure and a new memory-search task as the retrieval process measure. The two types of measure were not correlated with one another, but both correlated with span. Together they accounted for 40% of the variance in span.

Some of the difference between the proportion of variance accounted for by the model in the two experiments might reasonably be attributed to the additional variance in Experiment 1 stemming from age effects (absent, for all practical purposes, in Experiment 2). The raw correlation between age and span in Experiment 1 was .43, indicating that it could account for 18% of the total variance in span. Given that 87% of this age-related variance was accounted for by our model, it is simple to calculate that the 60% of the span variance accounted for overall included only 54% of the non-age-related variance, that is, (.87 × .18) + (.54 × .82) = .60. The difference between this 54% figure and the 40% of the span variance accounted for in Experiment 2 could be attributed to many factors, including but not limited to (a) the use of potentially more precise measures in Experiment 1 and (b) the variance added by the use of both letter and digit stimuli in Experiment 2.

This research takes a new tack and there are, not surprisingly, many unknowns remaining. For example, it is not yet certain to what extent the pauses within participants' repetition of the list in a span task (measured in Experiment 1) include the same processes as memory search, even though the suggested interpretation was that they do have processes in common. Providing a tentative initial answer to this question, Hulme, Newton, and Cowan (1997) obtained correlations between the slope of a search task modeled directly after Sternberg (1966) and interword pauses within correct responses in a span task administered to 24 adults. For monosyllabic word stimuli, the correlations between the slope in a search task and mean interword pauses for lists of Length 3–6 were $r =$ .20, .30, .31, and .58, respectively (though only the last correlation reached significance). In the present study, the durations of interword pauses, like search slopes in a novel type of search task, were correlated with span but not with rehearsal-related measures. Other reasons to suggest that interword pauses require some type of short-term memory search include the absence of word-length effects on either the interword pauses in memory-

span responses or the search times in Sternberg's probe task (Chase, 1977; Clifton & Tash, 1973; Cowan et al., 1994) and the dependence of both types of these types of timing measure on the length of the list held in memory (see Table 1; see also Cowan, 1992; Cowan et al., 1994; Sternberg, 1966). At any rate, span clearly involves some memory-retrieval process that is distinct from processes described in the phonological loop mechanism of working memory.

Reevaluating Global Rate-of-Processing Models

Our findings suggest that the view of individual differences in information processing that emphasizes a single, global rate of processing (e.g., Hale & Jansen, 1994; Kail & Salthouse, 1994) is an oversimplification. At least two separate types of processing rate are important. Similarly, Reed and Jensen (1993) found that choice reaction time and visual nerve conduction velocity correlated with intelligence, but not with each other. Such results suggest that there are important individual differences in the rate of processing beyond global rate and beyond any single specific mechanism such as rehearsal rate.

We hasten to add that our results do not rule out the existence of a general processing rate that would be analogous to the g factor in intelligence. We just would prefer not to call it a global rate given that more specific rates also exist. Such a general rate factor might, however, be what measures of global processing rate have picked up. For example, imagine a modification of Figure 1 in which the only path from age to another latent variable was to a general processing rate factor, which itself would have paths to the retrieval and rehearsal factors (as age does in our model). Imagine further that the general rate factor is measured in the same way that global processing rate has been measured and that the paths from this general rate factor to retrieval and rehearsal rates remain significant. That model would indicate that there exist both a general processing rate and more specific rate subcomponents (retrieval and rehearsal) that are not correlated with one another even though both are influenced by the general rate of processing. The viability of this hypothetical model shows that there is not necessarily an empirical discrepancy between the evidence underlying previous global rate-of-processing models and our evidence. Our factors may well *supplement* a general processing rate rather than replacing it.

Differentiating Age Effects Versus Individual Differences

In Experiment 1, the differences between measures that correlated with span and those that correlated with age also provide important information about the nature of cognitive development. Most notably, measures of the preparatory intervals between the stimulus list and the response in both speeded and unspeeded tasks consistently failed to correlate with span, even though some of these measures correlated with age (Appendix B). The most striking example may be the preparatory intervals for responses to two-digit lists in the span task. These intervals were uncorrelated ($r = -.06$)

with span in that same task, but still were highly correlated ($r = -.47$) with age. Other measures (rapid list-articulation durations for List Lengths 2 and 3) correlated with span ($rs = .37$ and .39), but not age ($rs = -.03$ and $-.15$), as shown in Table 2.

The pattern of correlations suggests that there are stable individual differences in span that are not accounted for by development throughout a wide range of ages (Grades 1–5). Such individual differences are noteworthy because they support arguments against a purely maturational model of intelligence. Typically, intelligence in children is measured in terms of "maturational age," the theoretical assumption being that a younger child who is relatively advanced for his or her age displays mental functioning that is equivalent to an older child who is relatively less advanced. At least in the case of memory span, the suggestion here is that there may be individual differences in the quality of mental functioning that are not closely associated with mental development.

Updating the Model of Memory Span

It is not yet clear what theory can account for the data from our study. The working memory model as proposed by Baddeley (1986) cannot account for it because it attributes memory-span differences to one factor, the rate of rehearsal. However, a modified account could be constructed from two of the basic mechanisms of that model, the phonological loop and the central executive. On one hand, the rate of rehearsal or articulation may serve as an index of how efficiently the phonological loop operates, that is, the explanation of span differences that Baddeley suggested. On the other hand, a rote rehearsal strategy is not always used in the span task. When other strategies are used, they may depend heavily upon the efficiency of central executive functioning. The rate of retrieval in the presence of a memory load, which presumably was indexed by the retrieval measures we used (interword pause durations in Experiment 1 and memory-search slopes in Experiment 2), may reflect the efficiency of central executive functioning.

The findings of Logie et al. (1996) support the suggestion that the function of a mental faculty other than the phonological loop is important in verbal memory span. Logie et al. examined individual differences in span and found that a variety of strategies were reported, including rehearsal, semantic coding, chunking of the list into smaller groups of items, retention of the first letter of each item, visual recoding, and mixed strategies. The magnitudes of effects that were taken to be indicative of the use of a phonological loop (phonological similarity and word-length effects) were higher when rehearsal or chunking strategies were reported than when other strategies were reported. These other strategies would require repeated planning and management that is usually attributed to the central executive component of the model. Similarly, Hulme, Maughan, and Brown (1991) found that the familiarity of items contributed to memory span in a way that was separate from the contribution of the time it took to articulate the list, whereas in the original working memory model, familiarity should be

irrelevant and only articulation time should matter. In terms of the working memory model, this suggests that participants often can engage in multiple forms of list encoding and retention, using both the phonological loop and central executive processes together.

If it turns out that the theoretical account of our study is accurate and that two critical factors determining an individual's memory span are rehearsal efficiency and short-term memory-retrieval efficiency, more work will be needed in order to determine why these factors are critical. One account (Cowan, 1992; Cowan et al., 1992, 1994) would state that the rates of covert rehearsal and retrieval from short-term memory are both important because these processes both must be carried out before a certain amount of information has been lost from short-term memory storage.

Remaining Issues Regarding the Mechanisms of Memory Span

Questions About the Role of Rehearsal

In the recent literature on memory span, there has been some question about the role of rehearsal. Cowan et al. (1992) found large effects of word length that resulted from output duration differences rather than rehearsal differences. Brown and Hulme (1995) showed that many of the effects of word length could be modeled using only decay or interference rather than rehearsal. However, there are reasons to continue to believe that rehearsal plays an important role in memory span. In a developmental study of immediate memory, Henry (1991) eliminated output effects by using a probe response, and she found that this method of testing eliminated word-length effects in 5-year-olds but not in 7-year-olds. Consistent with the previous literature on rehearsal processes in children (e.g., Flavell, Beach, & Chinsky, 1966), it seems reasonable to believe that this age difference resulted from rehearsal processes taking place in the older children but not the younger children. There is other research indicating that the effort required for rehearsal decreases across ages in childhood (Guttentag, 1984), so there is good reason to continue to suspect that the efficiency or rate of rehearsal accounts for the linear relation observed between articulation rate and memory span across different ages.

The doubt expressed regarding the role of rehearsal in immediate memory has been focused primarily on the relation between rehearsal and the word-length effect. Consider that Cowan et al. (1994) found that age and word-length effects come from different sources, with age affecting the time between words in the response as well as the total duration of the response and word length affecting the duration of words in the response, but not the time between words or the total duration of the response. It is quite possible, therefore, that rehearsal is not needed to explain the word-length effect, just as Brown and Hulme (1995) suggested. It still may be essential, however, to explain some other aspects of span performance, including the effect of age. Our results (e.g., Figures 1 and 2) appear to reconfirm that rehearsal plays a major role in recall.

Why might one expect rehearsal not to be an important determinant of the word-length effect? In addition to the arguments offered by Brown and Hulme (1995), perhaps rehearsal does not take place in a simple, repeating loop as Baddeley's (1986) theory suggested. Adult participants might find some way to reduce the impact of word length in their rehearsals of real words, such as by rehearsing only the beginning of each long word. This suggestion yields the prediction that rehearsal will prove to be a more important factor in producing the word-length effect for nonwords than for words, though presently there is little relevant evidence.

The Potential Role of Proactive Interference

Our interpretation might be questioned by some investigators on the grounds of very recent findings relating working memory to proactive interference. May, Kane, Hasher, and Valenti (1996) tested younger adults and older adults using a working memory-span procedure (modeled after Daneman & Carpenter, 1980) in which participants were to comprehend sentences and also retain the last word of each sentence. The working memory span was defined as the longest list of sentence-final words that could be retained. When this type of span was obtained by starting at the shortest list length and gradually increasing the list length (the usual procedure), the younger adults performed at a much higher level than the older adults. However, when a descending order of list lengths was used, this age difference disappeared. The interpretation offered was that there was proactive interference from previous trials at the longer list lengths when those longer lengths were presented last and that age differences in span resulted from differences in the susceptibility to proactive interference. It is possible that the same interpretation could be offered for age differences and individual differences in ordinary span within our study, inasmuch as list lengths within the span task were presented in an ascending order rather than a random, counterbalanced, or descending order.

We would respond to this point in several ways. First, there is a wealth of literature on span measured with ascending list lengths (as is used, for example, within tests of intelligence) and very little on span measured with a counterbalanced or randomized order of list lengths. Therefore, had we decided to use a counterbalanced or random order, our results would not have helped much in the interpretation of the past literature.

Second, using a counterbalanced or random order of list lengths would not eliminate proactive interference, although it might reduce it. Moreover, it would add variability to the span measurements that would lower their reliability. One reason that extra variability might obtain is that longer list lengths presented early on, when not much practice has been acquired, might prove especially overwhelming to some of the younger children. Use of a descending order, of course, does not seem well-motivated in the long run, though it was useful for May et al. (1996) to demonstrate the role of proactive interference.

Third, even if proactive interference does prove to be involved in the age difference in span, that does not

necessarily mean that the articulatory loop type of theory (or our modification of that theory) is inappropriate. Instead, it could be that the maintenance of phonological information in memory is useful precisely as a weapon against excessive proactive interference. Indeed, Tehan and Humphreys (1995) obtained results suggesting that transient phonemic codes provide immunity to proactive interference. In the critical trials within some of their experiments, participants received two 4-item blocks of items and were to recall only the second block. Proactive interference came from a semantic similarity between the items in the first and second blocks on some trials. An effect of proactive interference was obtained in circumstances in which the phonetic code was presumed to be weak (e.g., after 2 s of verbal shadowing), but not when the phonetic code was presumed to be strong (e.g., in immediate recall). Given that the phonetic code has this effect, span differences between individuals could be attributed either to differences in the amount of proactive interference or to differences in the ability to use the articulatory loop to overcome that interference.

A few other studies have examined the effects of proactive interference on the recall of lists of longer lengths. Halford, Maybery, and Bain (1988) used a memory-search procedure and found an effect of proactive interference on both the reaction time and the percentage correct with set sizes of 6, 8, and 10, but not with a set size of 4. In 8- and 9-year-old children, proactive interference was obtained for a set size of 4, but not a set size of 2. These results suggest that proactive interference emerges only when the set size is too large for all items to be maintained in an active form at the same time (e.g., perhaps within the focus of attention), requiring that some items be retrieved from less immediately accessible sources of memory. Thus, if proactive interference is involved in memory span, that could occur because participants with different spans differ in the set size at which the items can no longer be held in an active form. In turn, this could occur either because of a capacity difference (e.g., in the size of the focus of attention) or because of a difference in the ability to use strategies such as rehearsal to keep the phonological representation active. Either type of difference could result in differences in how much proactive interference is obtained at a particular list length. This appears to be fertile ground for future research.

To the extent that proactive interference is a source of individual differences and age effects in memory span, our results may be specific to the type of procedure in which a small set of highly familiar items is repeated over and over and correct serial order is required, all of which maximizes the chances of proactive interference. It may be only for such a procedure that the rehearsal process is needed to reactivate the phonological representation enough to counteract interference. This is the case, though, in the ordinary memory-span task used in tests of aptitude and intelligence.

A finding of LaPointe and Engle (1990) tends to support the speculation that similar results would not be obtained with a large set of stimuli. When they used a different set of words on each trial in a span task, word-length effects were still obtained, but the word-length effects were not eliminated by articulatory suppression as they were when a small

set of items was used over and over. This pattern of results can be accounted for in an articulatory loop type of model, once it is considered that participants were to read each word silently as it was presented. With a small set of stimuli used over and over, all of the stimulus words are in some sense active in memory, but this does not help recall. Rehearsal processes might be essential to reactivate and link words in the order that they appear in the current list, and without it a phonological representation may not be useful in recall. Thus, the word-length effect disappears with articulatory suppression. In contrast, when a new word is presented on each trial, its phonological representation might often be able to move from an inactive state to an active state and be linked to other items in the list, even in the presence of articulatory suppression (for supporting evidence see Besner, 1987). Once the representation is active, output effects or other decay-like effects on the phonological representation can produce the word-length effect.

The Role of Item Identification

There are other findings that also need to be reconciled with our theoretical view. Like our Experiment 1, Hitch, Halliday, and Littler (1989, 1993) examined two timing factors in the development of memory span in children. One was the rate of articulation. However, unlike our study, the other was the rate at which items could be identified, measured as the time to begin repeating a spoken word. These studies showed that identification times were correlated with articulatory times, but they accounted for individual differences in span only under circumstances in which articulation was suppressed. Although it was suggested that identification time depended on central executive processes, it might actually depend on both those processes and the phonological loop. There was no clean dissociation between types of timing measure as was obtained in our study.

The Role of the Central Executive

Finally, more work is needed to determine if the short-term memory-retrieval process really should be thought of as an index of central executive processing per se.[1] Other factors that seem promising as indexes of central executive processing in working memory include the ability to inhibit

---

[1] Another factor, the contribution of passive short-term memory persistence to memory-retrieval ability, also must be considered. Short-term memory of an item presented acoustically appears to decay over time at a rate that differs among adult individuals (Lü, Williamson, & Kaufman, 1992) and slows down with age in childhood (T. A. Keller & Cowan, 1994; Saults & Cowan, 1996). A more intact memory representation could make the work of the retrieval mechanism easier and thus could be the basis of faster short-term memory retrieval in some individuals than in others. Arguing against that interpretation, however, it might be expected that the efficiency of rehearsal, like retrieval, would depend on the integrity of the phonological short-term memory representation, in which case rehearsal and retrieval should have been correlated with one another. We found no such correlation.

irrelevant information (e.g., Conway & Engle, 1994; Gerns-bacher, 1993; Hasher & Zacks, 1988; Sullivan, Faust, & Balota, 1995), the ability to generate random numbers (Baddeley, 1986; Teasdale et al., 1995), and the ability to carry out a dual task (Baddeley, 1996; Baddeley & Hitch, 1974). If short-term memory retrieval heavily taxes central executive processes, the rate of short-term memory retrieval should correlate with these measures.

Conclusion

We have examined two types of timing factor that correlate with memory span but not with one another. They cannot both simply be indices of a common, global rate of processing. There might be a general rate measure that would correlate with both measures, though there would also be separate variances unique to one or both of the two rates we have examined. Perhaps only the measures that tap central executive processing reflect a general processing rate, whereas the efficiency of articulatory processing is more separate from this general rate. This type of study may thus prove relevant to the current debate over whether there are language-processing modules that remain separate from general working memory resources (Cantor, Engle, & Hamilton, 1991; Caplan & Waters, 1995; Martin, 1995; Miyake, Carpenter, & Just, 1995). Finally, we show that the factors significantly related to memory span are not always the same as the factors that change with age in childhood, relevant to the issue that Engle (1996) raised. Investigations of developmental and individual differences like our study can help to integrate cognitive and psychometric approaches to the study of mental functioning.

References

Baddeley, A. (1986). *Working memory*. Oxford, England: Clarendon Press.

Baddeley, A. (1992, January). Working memory. *Science, 255*, 556–559.

Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5–28.

Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

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

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology, 39A*, 467–478.

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

Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence, 15*, 229–246.

Caplan, D., & Waters, G. S. (1995). Aphasic disorders of syntactic comprehension and working memory capacity. *Cognitive Neuropsychology, 12*, 637–649.

Cavanagh, J. P. (1972). Relation between the immediate memory span and the memory search rate. *Psychological Review, 79*, 525–530.

Chase, W. G. (1977). Does memory scanning involve implicit speech? In S. Dornic (Ed.), *Attention and performance VI: Proceedings of the Sixth International Symposium on Attention and Performance* (pp. 607–628). Hillsdale, NJ: Erlbaum.

Clifton, C., & Tash, J. (1973). Effect of syllabic word length on memory-search rate. *Journal of Experimental Psychology, 99*, 231–235.

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

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*, 163–191.

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

Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition, 21*, 162–167.

Cowan, N. (1994). Mechanisms of verbal short-term memory. *Current Directions in Psychological Science, 3*, 185–189.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series, No. 26. New York: Oxford 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., 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 & Language, 33*, 234–250.

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

Deary, I. J. (1995). Auditory inspection time and intelligence: What is the direction of causation? *Developmental Psychology, 31*, 237–250.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Engle, R. W. (1996). Working memory and retrieval: An inhibition-resource approach. In J. T. E. Richardson, R. W. Engle, L. Hasher, R. H. Logie, E. R. Stoltzfus, & R. T. Zacks (Eds.), *Working memory and human cognition* (pp. 89–119). New York: Oxford University Press.

Flavell, J. H., Beach, D. H., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283–299.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*, 237–241.

Gathercole, S. E., Adams, A.-M., & Hitch, G. J. (1994). Do young children rehearse? An individual-differences analysis. *Memory & Cognition, 22*, 201–207.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove, United Kingdom: Erlbaum.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science, 4*, 294–298.

Gillam, R. B., Cowan, N., & Day, L. S. (1995). Sequential memory in children with and without language impairment. *Journal of Speech and Hearing Research, 38*, 393–402.

Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, 179–203.

Guttentag, R. E. (1984). The mental effort requirement of cumulative rehearsal: A developmental study. *Journal of Experimental Child Psychology, 37*, 92–106.

Hale, S., & Jansen, J. (1994). Global processing-time coefficients characterize individual and group differences in cognitive speed. *Psychological Science, 5,* 384–389.

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

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

Henry, L. A. (1991). The effects of word length and phonemic similarity in young children's short-term memory. *Quarterly Journal of Experimental Psychology, 43A,* 35–52.

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

Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 1–15). Thousand Oaks, CA: Sage.

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. (1997). *The nature of length and lexicality effects on short-term memory: A response timing analysis.* Manuscript in preparation.

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., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology, 47,* 72–87.

Kail, R., & Park, Y.-S. (1994). Processing time, articulation time, and memory span. *Journal of Experimental Child Psychology, 57,* 281–291.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica, 86,* 199–255.

Keller, T. A., & Cowan, N. (1994). Developmental increase in the duration of memory for tone pitch. *Developmental Psychology, 30,* 855–863.

Landauer, T. K. (1962). Rate of implicit speech. *Perceptual & Motor Skills, 15,* 646.

LaPointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1118–1133.

Loehlin, J. C. (1992). *Latent variable models* (2nd ed.). Hillsdale, NJ: Erlbaum.

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory & Cognition, 24,* 305–321.

Lü, Z.-L., Williamson, S. J., & Kaufman, L. (1992, December). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science, 258,* 1668–1670.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391–410.

Martin, R. C. (1995). Working memory doesn't work: A critique of Miyake et al.'s capacity theory of aphasic comprehension deficits. *Cognitive Neuropsychology, 12,* 623–636.

May, C. P., Kane, M. J., Hasher, L., & Valenti, M. (1996, November). *Proactive interference and working memory span.* Paper presented at the annual meeting of the Psychonomic Society, Chicago.

McGarry-Roberts, P. A., Stelmack, R. M., & Campbell, K. B. (1992). Intelligence, reaction time, and event-related potentials. *Intelligence, 16,* 289–314.

Miyake, A., Carpenter, P. A., & Just, M. A. (1995). Reduced resources and specific impairments in normal and aphasic sentence comprehension. *Cognitive Neuropsychology, 12,* 651–679.

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

Puckett, J. M., & Kausler, D. H. (1984). Individual differences and models of memory span: A role for memory search rate? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 72–82.

Reed, T. E., & Jensen, A. R. (1993). Choice reaction time and visual pathway nerve conduction velocity both correlate with intelligence but appear not to correlate with each other: Implications for information processing. *Intelligence, 17,* 191–203.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Salthouse, T. A., & Coon, V. E. (1993). Influence of task-specific processing speed on age differences in memory. *Journal of Gerontology: Psychological Sciences, 48,* P245–P255.

SAS Institute, Inc. (1989). *SAS/Stat user's guide, Version 6* (4th ed., Vol. 1). Cary, NC: Author.

Saults, J. S., & Cowan, N. (1996). The development of memory for ignored speech. *Journal of Experimental Child Psychology, 63,* 239–261.

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., Guentert, L., & Hersberger, L. (1990). Phonological similarity, pronunciation rate, and memory span. *Psychological Science, 1,* 74–77.

Sipe, S., & Engle, R. W. (1986). Echoic memory processes in good and poor readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 402–412.

Spearman, C. (1927). *The abilities of man.* London: MacMillan.

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

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time processes. *American Scientist, 57,* 421–457.

Sternberg, S. (1975). Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology, 27,* 1–32.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 469–505). New York: Academic Press.

Sullivan, M. P., Faust, M. E., & Balota, D. A. (1995). Identity negative priming in older adults and individuals with dementia of the Alzheimer type. *Neuropsychology, 9,* 537–555.

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition, 23,* 551–559.

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

Verive, J. M., & McDaniel, M. A. (1996). Short-term memory tests in personnel selection: Low adverse impact and high validity. *Intelligence, 23,* 15–32.

Vernon, P. A., & Mori, M. (1992). Intelligence, reaction times, and peripheral nerve conduction velocity. *Intelligence, 16,* 273–288.

Zhang, G., & Simon, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. *Memory & Cognition, 13,* 193–201.

Appendix A

Results for Auxiliary Timing Measures in Experiment 1

Preparatory Intervals in the Span Task

Preparatory intervals in the span-task responses may be more complex than interword pauses, reflecting not only retrieval processes on a local basis but also additional processes, such as articulatory planning relevant to the entire response. Thus, these periods typically are much longer than interword pauses. Although they did not produce a significant age effect, there was a significant effect of list length, $F(2, 138) = 32.69$, $MSE = 16,619$, $p < .001$, and an Age Group $\times$ List Length interaction, $F(4, 138) = 3.37$, $MSE = 16,619$, $p < .02$. The mean preparatory intervals for responses to two-word lists were 894, 813, and 698 ms in the first, third, and fifth graders, respectively. For three-word lists, the comparable durations were 746, 684, and 656 ms, and for four-word lists, the durations were 628, 642, and 618 ms. The age group effect was significant only for two-word lists. This pattern of preparatory intervals is interesting because it differs from speeded responses, for which longer lists produce longer rather than shorter periods of silent preparation (see Sternberg et al., 1978, as well as our speeded articulation task). The reason for this difference between tasks is unclear. Perhaps, inasmuch as participants have as long to respond as they would like in this task, longer lists elicit responses more promptly because participants realize that they are in danger of forgetting the material quickly over time. At any rate, the results mesh with an assumption stemming from some of the previous research (Cowan, 1992) that preparatory intervals are not as easily interpretable as interword pauses. Thus, they are not included in our latent variable model of span.

Word Durations in the Span Task

A second set of auxiliary measures comprises durations of spoken words (i.e., digits) within the span-task responses. These durations increased as the list length increased, $F(2, 138) = 12.89$, $MSE = 2,139$, $p < .001$. However, they decreased as a function of age, $F(21, 69) = 10.61$, $MSE = 35,650$, $p < .001$. These two factors did not interact. For two-item lists, the mean digit durations were 540, 485, and 411 ms in the three age groups, respectively; for three-item lists, the comparable durations were 571, 517, and 414 ms; and for the four-item lists, the durations were 580, 537, and 437 ms. The digit durations are not clearly interpretable because they could reflect either the intrinsic rate of articulation processes or the influence of concurrent search processes or load effects on articulation. The effect of list length on digit duration suggests that such concurrent processing does take place. Given this complexity and the poor correlation between spans and word durations in the span task previously (Cowan, 1992; Cowan et al., 1994), the digit durations, like the preparatory intervals, are not included in our latent variable model of span.

Serial-Position Effects in Span-Task Responses

A more refined understanding of digit durations and pauses in the span task might be obtained by examining them separately for each serial position, yielding clues about how the list is processed. If the list responses were to conform to the model set out by Sternberg et al. (1978) for speeded responses, there should be no serial-position effects because an exhaustive search is said to take place before the production of each word. However, there were such serial-position effects. In the analyses of interword pauses, three-digit lists yielded a significant serial-position effect, $F(1, 69) = 9.97$, $MSE = 2,080$, $p < .003$. The mean pause was 205 ms following the first digit in the response, but only 181 ms following the second digit. Similarly, in four-digit lists, there was again an effect of serial position, $F(2, 138) = 4.83$, $MSE = 3,282$, $p < .01$; the mean pauses were 260 ms, 256 ms, and 233 ms, with the last pause differing from the first two, $ps < .05$. There was no significant effect of serial position in digit-duration analyses for two- or three-digit lists, but there was such an effect for four-digit lists, $F(3, 207) = 16.28$, $MSE = 4,411$, $p < .001$. The mean durations of digits in Serial Positions 1–4 of the responses were 519 ms, 557 ms, 514 ms, and 480 ms, respectively. In Newman-Keuls pairwise comparisons, the first and third serial positions did not differ but all other pairs of means differed at the $p < .01$ level.

Although the bases of these serial-position effects are unclear, one can consider the potential contribution of several factors. First, it is possible that items that have just been pronounced are in some way marked in short-term memory. (For a similar concept applied to sentence production, see Dell, 1986). When all items from the list but one have been pronounced, the search process to identify the item to be pronounced last may be facilitated because all items but one are marked. Consistent with this logic, the last interword pause and the last word duration tended to be smaller than previous ones in the response. (This also could occur because the list has been learned better by the time of the last pause.) Second, there is a potential effect of prosody or grouping, which could explain the unusually long duration of the second word in responses to four-word lists. These effects are documented here to inspire future research, but they will not be discussed further in this article.

Preparatory Intervals in the Rapid List-Articulation Task

Preparatory intervals increased across list lengths, $F(3, 207) = 27.64$, $MSE = 43,499$, $p < .001$, suggesting that children may in some manner process (e.g., plan) the entire list pronunciation during these intervals, similar to what Sternberg et al. (1978) found in an adult study. For List Lengths 1–4 (across trials) the mean preparatory intervals were 440, 487, 511, and 567 ms, respectively.

There were large main effects also of repetition, $F(5, 345) = 24.67$, $MSE = 31,294, p < .001$. On Trials 1–6 for a particular list length, collapsed across list lengths, the mean preparatory intervals were 603, 493, 499, 479, 473, and 461 ms, indicating improvement with practice. These factors did not interact. Finally, preparatory intervals in this task decreased with age, $F(2, 69) = 5.71$, $MSE = 303,069, p < .01$. The mean intervals were 558 ms in first graders, 497 ms in third graders, and 449 ms in fifth graders. This age effect did not interact with either list length or repetition, $F < 1$. Given the uncertainty about the interpretation of preparatory intervals (Cowan, 1992), this measure was not included in the latent variable model.

Preparatory Intervals in the Rapid Counting Task

In the analysis of preparatory intervals in the task of counting from 1 to 10, the age effect fell short of significance. There was an effect of repetition, $F(3, 207) = 7.49, MSE = 78,242, p < .001$. For the four repetitions the means were 810 ms, 622 ms, 734 ms, and 629 ms, respectively. The first and third repetitions differed from the second and fourth, $ps < .05$. This effect of repetition, which did not interact with age, is not easily explained. Like the other preparatory intervals, this one was not entered into the latent variable model.

Appendix B

Correlations Involving Auxiliary Timing Measures in Experiment 1

Measure	Pr2	Pr3	Pr4	Wd2	Wd3	Wd4	Rp1	Rp2	Rp3	Rp4	Cp
Age	−.47**	−.26	−.08	−.48**	−.47**	−.45**	−.38**	−.34**	−.39**	−.35**	−.19
Mean span (MSp)	−.06	−.01	.16	−.27	−.32**	−.43**	−.11	−.11	−.21	−.05	.09
Span Test 1 (Sp1)	−.02	.03	.20	−.20	−.27	−.37**	−.06	−.10	−.11	.04	.15
Span Test 2 (Sp2)	−.09	−.05	.09	−.30**	−.33**	−.42**	−.15	−.11	−.28**	−.13	.01
Preparatory intervals in the memory-span task (Pr)											
Length 2 (Pr2)	—										
Length 3 (Pr3)	.42**	—									
Length 4 (Pr4)	.39**	.61**	—								
Word durations in the memory-span task (Wd)											
Length 2 (Wd2)	.48**	.46**	.20	—							
Length 3 (Wd3)	.47**	.45**	.21	.87**	—						
Length 4 (Wd4)	.39**	.43**	.15	.83**	.93**	—					
Interword pauses in the memory-span task (Ps)											
Length 2 (Ps2)	.31**	.33**	.19	.62**	.63**	.61**	−.05	.07	.17	.02	.06
Length 3 (Ps3)	.32**	.42**	.24	.63**	.70**	.67**	−.06	.00	.11	.01	.08
Length 4 (Ps4)	.30**	.27	.17	.53**	.64**	.61**	.00	.12	.08	.04	.10
Preparatory intervals in the rapid list-articulation task (Rp)											
Length 1 (Rp1)	.26	.20	.24	.05	.05	.09	—				
Length 2 (Rp2)	.26	.29**	.26	.16	.18	.19	.75**	—			
Length 3 (Rp3)	.29**	.38**	.35**	.18	.24	.22	.55**	.62**	—		
Length 4 (Rp4)	.25	.39**	.39	.15	.17	.19	.61**	.71**	.60**	—	
Speaking durations in the rapid list-articulation task (Ra)											
Length 1 (Ra1)	.24	.11	−.01	.17	.20	.24	.04	−.15	−.14	.04	.01
Length 2 (Ra2)	.02	−.12	−.21	−.07	.03	.11	.15	−.03	−.09	.04	−.03
Length 3 (Ra3)	.04	.06	−.21	.00	.02	.16	.20	.06	.00	.05	.04
Length 4 (Ra4)	.03	−.03	−.34**	.13	.18	.30**	.13	.06	.02	.03	.01
Preparatory intervals in the rapid counting task (Cp)											
Cp	.22	.41**	.25	.11	.10	.09	.52**	.52**	.47**	.53**	—
Speaking duration in the rapid counting task (Cd)											
Cd	.04	.09	−.08	.07	.06	.15	.35**	.33**	.21	.38**	.35**

$**p < .01$, one-tailed.

(Appendixes continue)

Appendix C

Comparison of Search Tasks

We conducted a smaller, supplementary study with 32 new adult participants to estimate the relationship between the array search task of Experiment 2 and the traditional search task developed by Sternberg (1966, 1969). In Task 1, each participant carried out a computer-administered version of Sternberg's task (1966) with one- to six-digit stimuli in the memory set on each trial. There were 4 practice trials. On positive test trials, there were 2 trials with a probe corresponding to each serial position for each set size. There also was an equal number of negative trials for each set size, making 84 trials in all, and the trial order was randomized. After each response to a probe, the participant was to recall the entire positive set by making keypress responses. The probe task yielded a mean positive slope of 35.72 ms/item ($SD = 53.20$) and a mean negative slope of 77.06 ms/item ($SD = 59.08$), suggesting a relatively slow, self-terminating search. Task 2 comprised letter- and digit-array searches identical to those used in Experiment 2. This yielded much higher slopes similar to our Experiment 2: for letters, 239.57 ms/item ($SD = 141.65$), and for digits, 174.35 ms/item ($SD = 75.67$). Finally, Task 3 comprised another task like that of Sternberg, with memory sets of one through four, an equal number of positive and negative trials, and 18 trials for each

memory set size, all randomized. This task yielded a mean positive slope of 37.10 ms/item ($SD = 90.85$) and a mean negative slope of 22.56 ms/item ($SD = 55.03$), suggesting a rapid and exhaustive search.

Correlations were calculated across averaged slopes to increase the stability of the measures. The correlation between the average slopes of Tasks 1 and Task 3 was low and negative in sign, $r = -.18$. The Task 2 (array search) slope, averaged across letter and digit trials, when correlated with the mean slope of Task 1 produced $r = .39, p < .05$, two-tailed, but the correlation between Task 2 and Task 3 was minuscule, $r = .01$. The results suggest that the search task used in Experiment 2 was likely to involve self-terminating searches, like Task 1 but unlike Task 3 or the typical tasks modeled after Sternberg (1966). This slower, self-terminating search seems more reasonable for the type of processing that would occur during memory-span task performance.