

Short-Term Memory, Working Memory, and Verbal Abilities: How Do They Relate?

JUDY CANTOR,
RANDALL W. ENGLE,
AND GEORGE HAMILTON
University of South Carolina

Subjects performance on short-term memory (STM) spans, STM probe-recall tasks, and complex working memory (WM) spans was used to assess the relationship between STM and WM, and to test whether these measures independently relate to verbal abilities. Factor analysis indicated that scores on the STM spans and probe-recall tasks loaded on a factor that was distinct from the WM spans, and regression and part correlations showed that these different factors accounted for separate variance in the Verbal Scholastic Aptitude Test (VSAT). These results provided evidence that STM and WM are different cognitive constructs, both of which are important to verbal abilities. It was also shown that within STM measures, rehearsal can obscure the relationship between STM capacity and abilities. For example, in the probe-recall tasks, only performance on final list items correlated with VSAT, and it was argued that these items were the most recently represented in STM, but had the least opportunity to be rehearsed. It was also suggested that digits are easier to rehearse than words, and we found that a STM word span correlated with VSAT, but a STM digit span did not. Furthermore, in considering each serial position in the probe-recall tasks, fewer items correlated with verbal abilities when digits were used as stimuli than when words were used. These results converge on the notion that rehearsal drives down the correlation between STM capacity and abilities. Overall, it was concluded that STM and WM tasks do reflect different cognitive constructs, both of which seem to be important in abilities. In addition, when predicting verbal abilities from a STM tasks, the best measure is one in which rehearsal is not strongly influencing performance.

Many contemporary models of immediate memory distinguish between the simple storage properties of short-term memory (STM), and the more complex central processing functions of an active working memory (WM) (Anderson, 1983; Baddeley, 1986; Baddeley & Hitch, 1974; Daneman, 1987; Daneman & Carpenter, 1980, 1983). STM has traditionally been viewed as a temporary buffer, in which a limited number of items can be maintained using simple strategies such as rehearsal or chunking. By contrast, WM is often considered to

We wish to thank Lee Kirkpatrick for sharing his insights on statistical issues, and Gloria Miller, Ronald Cohen, and Robert Greene for their helpful comments on this manuscript.

Correspondence and requests for reprints should be sent to Judy Cantor, Department of Psychology, University of South Carolina, Columbia, SC 29208.

be a more flexible computational arena, with a limited amount of attentional resources that are required for the processing and manipulating of information, and for storing the partial and final products of these processes. The capacity of WM has been implicated in such complex behavior as problem solving and reasoning (Anderson, 1983; Baddeley, 1986; Baddeley & Hitch, 1974), reading comprehension (Carpenter & Just, 1988; Daneman & Carpenter, 1980, 1983; Engle, Nations, & Cantor, 1990; Turner & Engle, 1989), and even general intelligence (Daneman, 1982, 1987; Daneman and Tardiff, 1987; Larson & Alderton, 1988, 1990).

Baddeley and Hitch (1974) provided the first empirical work testing whether there is a unitary immediate memory system or distinct subsystems. They reasoned that if there were a single limited-capacity memory for all cognitive activity, then, by loading memory with a simple STM digit task, subjects should be prevented from completing the more complex tasks of reasoning or comprehension. If, on the other hand, there were separate memories for simple and complex processing, then maintaining digits should not result in reasoning or comprehension deficits. Interestingly, they found that loading memory with three digits did not affect concurrent task performance, but a six-digit load showed some impairment. Arguing that the capacity of STM is approximately three or four items (see also Broadbent, 1975), their findings implied that retaining small numbers of digits and performing more complex verbal tasks do not rely entirely on the same cognitive resources.

Baddeley and Hitch (1974) posited a tripartite memory with a central WM processor, and two slave systems. The central processor is the core of immediate memory, in which sophisticated processing and subsequent storage compete for limited attentional resources. It is the central resources of WM that are most important to complex cognitive activity. There are also systems subordinate to the central mechanism, one of which is an articulatory mechanism analogous to verbal STM. STM has its own limited capacity, and maintains items using a phonologically based rehearsal loop. Unless STM is disabled, or its capacity is exceeded, it can keep a limited amount of information active independently of central WM resources.

Baddeley and Hitch (1974) provided a model in which STM and WM seem to be separate constructs, and emphasized that the processing arena of working memory is particularly important to higher level cognition. With this premise, we might also expect that WM would be a more important factor in individual differences in cognitive abilities than would STM. Consistent with this notion, there are several reports in which differences in STM processes—such as encoding, search rate, rehearsal, or chunking—do not distinguish between high and low-ability subjects (see, for review, Cohen & Sandberg, 1980; Dempster, 1981; Perfetti & Lesgold). By contrast, measures of WM have consistently predicted scores on a variety of complex verbal tasks. For example, Daneman and Car-

penter (1980) developed a WM reading-span measure in which subjects read aloud a series of unrelated sentences, and then recalled the last word of each sentence. They argued that maintaining words while reading activated complex processes and demanded storage in WM, so that the number of final words that could be recalled in this task was a good indication of WM capacity. In the original work by Daneman and Carpenter (1980), and in several subsequent studies, measures of WM strongly predicted a variety of verbal abilities measures such as the Verbal Scholastic Aptitude Test (VSAT), standardized reading comprehension scores, the ability to recover from garden path passages, and to make inferences about novel words from context (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Masson & Miller, 1983; Turner & Engle, 1989). Moreover, the WM spans need not be reading related. When the sentences in the Daneman and Carpenter (1980) task are replaced with mathematical operations, WM spans still strongly correlate with verbal comprehension (Engle et al., 1990; LaPointe & Engle, 1990; Salthouse, Mitchell, Skovronek, & Babcock, 1989; Turner & Engle, 1989).

During the past few years, our own lab has concentrated on finding what is "special" about the WM tasks that enable them to distinguish high- and low-abilities subjects, and how these tasks differ from traditional STM spans. We have found, however, across several studies, that the STM and WM spans are surprisingly similar. Both, for example, are effected by word length and word knowledge (Engle et al., 1990; LaPointe & Engle, 1990). Although the WM spans are usually stronger predictors of abilities, STM and WM measures both significantly predict verbal comprehension (Engle et al., 1990; LaPointe & Engle, 1990) and are internally consistent (LaPointe & Engle, 1990).

We do not know why we generally find correlations between STM span and more complex tasks whereas others do not, but we can speculate on what we believe are the important differences between studies. Much of the earlier work on STM focused on processing characteristics such as encoding time, rehearsal, and search rate. Although these parameters may be important in STM, they do not seem to be good predictors of differences in abilities (see Cohen & Health, 1990; Cohen & Sandberg, 1977, 1980; Dempster, 1981; Perfetti & Lesgold, 1977). Our STM span is a more global capacity measure based on the number of items recalled. When only this type of capacity measure is considered, there is some positive evidence for a relationship between STM and abilities (Cohen & Sandberg, 1977, 1980; Das & Sui, 1989; Dempster & Cooney, 1982; Engle et al., 1990; LaPointe & Engle, 1990). Although there is also some negative evidence (Chiang & Atkinson, 1976; Daneman & Carpenter, 1980; Turner & Engle, 1989), the consistency with which we find a relationship leads us to argue that it is theoretically important and warrants further investigation. The data to be reported in this study extends the findings that STM capacity measures predict verbal abilities, and investigates alternative explanations of these results.

Two issues seem to stand out in recent studies that find STM-abilities relations. First, of course, is that measures of STM capacity predict verbal comprehension. Second, is the fact that scores on STM and WM tasks seem to intercorrelate. These findings have suggested to us that the complex WM spans and the simple STM spans measure the same construct (Engle et al., 1990; LaPointe & Engle, 1990). Following the models of Bower (1975) and Anderson (1983), WM was defined as the currently active portion of long-term memory knowledge (nodes, procedures, etc.) above some critical threshold, but below the level of conscious awareness (see LaPointe & Engle, 1990). WM is, in essence, whatever information is activated above resting state for current cognitive activity. Because the total amount of activation is limited, the capacity of WM will be manifest in most intelligent behavior. By contrast, STM is a simple limited-capacity buffer representing activated information that is attended at a superficial coding level for maintenance purposes only. Moreover, we thought that both STM and WM spans tasks require such superficial processing. What distinguished them was that the WM spans prevent the rehearsal and chunking strategies that reduce our ability to obtain a purer measure of the capacity of STM. The traditional STM spans allow for, and even encourage, the simple strategies like rehearsal that are unimportant to most complex information processing.

The notion that rehearsal can obscure the relationship between STM and abilities was also posited by Cohen and Sandberg (1980). To test this hypothesis, they used a STM probe-recall procedure in which trials of nine auditory digits were presented, followed by a probe to indicate that the first, middle, or final few items should be recalled. In work with children, they found that only performance on the last few items, and occasionally middle-list items correlated with IQ. They argued that these were the most recent in STM but have received the least opportunity for rehearsal. Recall of primacy items, those that should have been most rehearsed, did not predict intelligence scores (but, see Merkel & Hall, 1982). According to Cohen and Sandberg (1980), as the opportunity to rehearse increases, more items from the beginning of the list will be rehearsed, and performance on fewer serial positions will correlate with abilities. These results conform to our original notion that the best measure of STM capacity will be one in which rehearsal is least influential. These studies do not, however, address the issue of whether or not STM and WM spans reflect the same system.

The study to be reported here is part of an ongoing attempt to examine the relationship between STM and WM spans, and how rehearsal influences the relationship between memory capacity and abilities. Subjects were administered traditional STM spans, complex WM spans, and the probe-recall task of Cohen and Sandberg (1977, 1980). If the STM and WM spans differ only in the amount of rehearsal they encourage, and otherwise reflect the same cognitive system, then a clear pattern of correlations should emerge among performance on the

probe-recall, STM, and WM spans. Because the final few items of the probe-recall task are theoretically least affected by rehearsal, we would expect that performance on these items would most highly relate to the WM spans. Moreover, if the STM spans are simply poor indicators of the same capacity measured by the WM and probe-recall tasks, then the STM spans should not account for any additional variance in a measure of cognitive abilities when the influences of the other tasks are removed.

Finding that different span procedures intercorrelated, and predicted the same criterion measure provided initial evidence that they may also reflect the same construct. But it seemed to us that it is also possible to attribute the different tasks to different memories, both of which are important to cognitive abilities. Simply because two tasks predict the same criterion does not necessarily imply that they are a reflection of the same underlying mechanism. The WM span may still index the interplay between complex processing and storage in WM, whereas the STM spans measure the simple storage properties of a temporary buffer. The correlation between STM and WM tasks may reflect the shared characteristics of otherwise different systems, rather than a unitary memory. Thus, an alternative to the notion of a single immediate memory is that there are indeed separate memories that underlie the STM and WM spans. As argued by Baddeley and Hitch (1974), the cognitive system may have a dedicated storage arena that cannot be used for general processing. This STM system is responsible for maintaining small numbers of verbal items and places little demand on the central WM processor. The STM spans and the probe-recall measures are simple storage task and would, therefore, depend on such a system. On the other hand, the WM spans require processing and storage and should, therefore, be indicative of a more central WM processor. If, as posited by Baddeley and Hitch (1974), WM and STM are separate systems, then we would expect the probe-recall and STM tasks to intercorrelate, with neither relating strongly to the WM measures. Notice that, although this alternative hypothesis posits separate memories, it does not necessarily discount the notion that rehearsal obscures the relationship between STM tasks and abilities. It is still possible that within the STM system, tasks that are least influenced by rehearsal will be better measures of STM capacity and better predictors of cognitive abilities. However, if our measures reflect different underlying constructs, then we would expect that the STM span and probe-recall tasks would also capture different variance in a verbal abilities measure than the WM spans.

METHODS

Subjects

Subjects were 49 undergraduates from the University of South Carolina who received course credit for volunteering to participate and have their VSAT scores verified.

Materials

The probe-recall task has been previously studied only when the items to-be-remembered were single digits. The reading span, on the other hand, most often involves words. Because digits come from a limited pool and often occur together in the real world, they are probably easier to rehearse or chunk than unrelated words. Because we were concerned about the influence of these strategies on our measures, two versions of each task were generated, one with digits and the other with words.

All to-be-remembered words were high-frequency nouns, selected from the Francis and Kucera (1982) book of word norms. Separate pools were randomly generated for the STM span, WM span, and probe-recall task, with the constraint that no word appear more than once either within or between pools. For tasks involving digits, items were randomly generated so that no digits appeared more than once per trial, and so that the same pattern of digits was not presented across tasks.

The STM memory spans and the WM span were computer controlled, presented via IBM XT with an Amdek monochrome monitor. The probe-recall tasks were presented with a Sony stereo cassette recorder, with prerecorded stimuli by a male experimenter.

STM Span Tasks

On each trial of the STM digit and word-span tasks, stimuli were presented individually at a rate of one item per second, and subjects read aloud each item at presentation. The end of a trial was marked by the display of a single question mark, also used to cue the onset of recall. Subjects were permitted as much time as they needed to recall the lists. A new trial was initiated by a key press. List length ranged from three to nine items in the digit version of this task, and three to eight items for the word version. For each task, there were three trials at each list length, but the order presented was randomized so that length was not predictable.

WM Span Tasks

Sentences were the same as those used by LaPointe and Engle (1990), and were 13 to 16 words in length. For the WM digit and word-span tasks, a sentence was displayed on the screen above a memory item. Subjects read each sentence and then the corresponding memory item aloud, and then an experimenter pressed a key to display new stimuli or a question mark indicating the end of a trial and the onset of recall. For both the sentence-digit and sentence-word tasks, trial length varied from two to seven, with randomized presentation so that trial length was not predictable.

Probe-Recall Tasks

These tasks followed the same format as that of Cohen and Sandberg (1977, 1980). On each trial, nine items (digits or words) were presented at a rate of three items per second. Following the final item, and in synchrony with the list, was the probe A, B, C, or RECALL, indicating that subjects should write the first, middle, or last three items, or the entire list, respectively. There were 28 trials, 8 each for probing of A, B, and C, and 4 trials that corresponded to total list recall.

Procedure

Subjects participated individually in a single session of approximately 1.5 hr. Order of the three types of tasks (STM span, WM span, and probe-recall) was counterbalanced, but within each task type, the words were always presented prior to the digits. This produced six task orders. Unfortunately, due to technical problems, the number of subjects in each of these orders was only approximately equal (ranging from 6 to 9 subjects).

For all tasks, strict serial recall instructions were administered, and response booklets provided spaces for seven to nine items per trial (depending on the task) to be used for written recall. Three extra trials were presented at the beginning of each task to be used as practice, and performance on these trials was not used in any statistical analyses.

RESULTS

Scoring

For the STM-span and the WM reading-span tasks, a score was assigned that equalled the sum of the correctly recalled items in order, for those trials that were perfectly recalled. For the probe tasks, each serial position (1-9) was scored separately (excluding total recall trials) and then three measures were calculated, corresponding to the sum of the correct responses for the beginning, middle, or final three position. For those trials that required recall of the entire list, scores were assigned to serial positions 1-9, corresponding to the total number of items correctly recalled in each position.

Analyses

For all of our analyses, *p* values of .05 or better will be indicated by an asterisk. It should be noted that because task order was counterbalanced, there is a potential confounding of subject variance with variance due to order. However, we found no relationship between task order and any of our variables (all *r*s < .09). Table 1 shows the descriptive statistics for each of our measures. Prior to testing the notions of distinct versus a unitary memory system, it is necessary to show that our measures predict verbal abilities. Therefore, also provided in Table 1 are the correlations between our tasks and VSAT. Several points are noteworthy. The

TABLE 1
Descriptive Statistics and Correlations of Tasks
With VSAT

	Percent Correct	SD	Correlation with VSAT
WM Spans			
Words	15.3	9.41	.42*
Digits	33.7	14.25	.37*
STM Spans			
Words	30.2	12.30	.35*
Digits	40.4	19.55	.04
Probe-Recall			
Words A	47.7	3.74	-.07
B	20.0	4.28	.30*
C	67.0	5.17	.32*
Total Recall	16.1	3.32	.32*
Digits A	69.2	4.19	.11
B	41.6	6.19	.24*
C	74.3	5.25	.30*
Total Recall	38.6	6.54	.19

Note. $N = 49$; VSAT: $M = 433.5$, $SD = 99.0$.
* $p < .05$.

WM and STM word spans significantly predicted VSAT ($r = .42^*$, $.35^*$, respectively), and in the probe-recall word task, only performance on middle ($r = .30^*$) and final ($r = .32^*$) items correlated with verbal abilities. Initial list items, those expected to be most influenced by rehearsal, showed no predictive validity.

Similarly, the WM digit span correlated with VSAT ($r = .37^*$), as did the B and C items in the probe-recall digit task ($r = .24^*$, $.30^*$, respectively). All of these findings fit well with previously reported memory-span studies, and the work of Cohen and Sandberg (1977, 1980). Interestingly, however, the STM digit span did not predict verbal abilities ($r = .04$). We included both digits and words in this study because we thought that digits might be easier to rehearse, and that rehearsal may obscure a measure of "pure" STM capacity. We would then predict that tasks involving digits would correlate less with VSAT than those involving words. In the case of the WM digit span, the processing of sentences probably prevents rehearsal, so that it would be expected that performance on this task would relate to VSAT. Finding that the STM digit span doesn't predict abilities is also consistent with this notion. However, in examining the results of the probe-recall tasks, one might get the impression that rehearsal is not driving down the relationship between simple digit tasks and VSAT. Clearly, if digits are easier to rehearse, then we would expect more items to be rehearsed in simple digit tasks, and fewer corresponding serial position scores to correlate with

VSAT. As can be seen in Table 1, middle and final items in the probe-recall task correlated with VSAT, regardless of whether the stimuli were digits or words. But scores for A, B, and C in the probe-recall tasks are summations across three serial positions each. This does not provide very sensitive information about the number of serial position scores that actually predicted VSAT. It was, therefore, decided to break down each of the probe-recall tasks and compute the correlation between the scores at each serial position and VSAT. These results are presented in Table 2, and support the contention of Cohen and Sandberg (1980) that, as the opportunity to rehearse increases, more items from the beginning of the list will be rehearsed, and fewer serial positions will correlate with cognitive abilities.

Notice that, for the probe-recall word task, scores corresponding to positions 4–9 (B&C) relate to verbal abilities. This is true regardless of whether recall was of the whole list (total) or only of part of the list. For the digit task there are fewer overall serial position scores that correlate with VSAT, and those that do relate tend to be towards the end of the list, where rehearsal is least likely to influence performance.

The results of breaking down the probe-recall task by serial position indicates that there are only a few positions with scores that actually correlate with VSAT when digits are used. If we generalize these findings to the STM digit span, the proportion of trials in that task in which more than six serial positions were represented is less than half, so that we might expect that this task will not predict our criterion. Overall, these findings imply that, within STM measures, rehearsal obscures the relationship between STM capacity and complex cognitive abilities. Rehearsal may, for example, increase the functional capacity used in simple storage tasks, but is not indicative of the underlying capacity of STM in more

TABLE 2
Correlation Between Each Serial Position of the
Probe Tasks and VSAT

Serial Position	Words		Digits	
	Probe- Recall	Total Recall	Probe- Recall	Total Recall
1	.03	.01	.11	.18
2	-.04	.09	.10	-.07
3	.13	-.06	.08	.20
4	.33*	.28*	.18	.20
5	.32*	.27*	.27*	.15
6	.30*	.28*	.20	.16
7	.25*	.26*	.25*	.30*
8	.28*	.27*	.28*	.25*
9	.33*	.29*	.32*	.27*

Note. $N = 49$.
* $p < .05$.

complex situations when using resources for simple mnemonics is probably not fruitful.

We have suggested that rehearsal does influence the relationship between simple storage measures and abilities, but have not yet addressed whether or not our memory tasks reflect the same underlying construct. We argued that those tasks that measure the same memory should intercorrelate most highly. Table 3 represents the zero-order correlations among our measures. Scores from the first few serial positions of the probe tasks are omitted, because these correlated only with each other. Also, for these tasks, recalls of the middle and final items were summed, because they are considered to reflect the same mechanism regardless of which hypothesis is being tested. We acknowledge that including the B positions in the probe-recall digit task probably also increases the influence of rehearsal on the B + C score, but separating the B from C items did not change the pattern of results.

Notice that performance on the STM digit span, which is considered to be most indicative of rehearsal, did not predict the WM spans. These results are consistent with our previous conclusions about the role of rehearsal in the relationship between memory span and more complex behavior. One of the hypotheses to be tested is that all of our measures reflect the same system, distinguished only by the amount of rehearsal that they afford. We would then expect the B + C positions in the probe-recall tasks to be most highly correlated with the WM measures. This was not the case. The probe-recall word task did predict the WM spans, but the digit version of this task showed no relationship to the WM spans. Moreover, the probe-recall tasks were correlated with the STM measures, a finding contrary to expectations based on the notion that what distinguished STM and WM spans is the amount of rehearsal they allow.

An alternative to the notion that our tasks reflect the same mechanism, is that there are two memories that underlie performance. The simple memory tasks

TABLE 3
Intercorrelations Among Memory Measures and VSAT

	STM Spans		Probe B + C		WM Spans	
	Words	Digits	Words	Digits	Words	Digits
STM Words	—	—	—	—	—	—
STM Digits	.52*	—	—	—	—	—
Words B + C	.77*	.39*	—	—	—	—
Digits B + C	.36*	.48*	.74*	—	—	—
WM Words	.37*	.18	.30*	.20	—	—
WM Digits	.27*	.17	.24*	.11	.63*	—
VSAT	.35*	.04	.32*	.29*	.42*	.37*

Note. $N = 49$.

* $p < .05$.

(STM spans and probe-recall) may reflect STM, whereas the WM spans may involve the processing-storage tradeoff indicative of a separate WM. Based on this hypothesis, we would expect that the STM and probe-recall tasks would intercorrelate, but neither of these would strongly relate to the WM measures. As seen in Table 3, the simple memory tasks do intercorrelate, but the relationship between these measures and the WM spans is more complicated. The simple memory measures involving words predict the WM spans, whereas the simple memory measures involving digits do not. Overall, although a model of separate memories seems to describe the data better, finding that the STM word span and probe-recall word task correlate with the WM spans is contrary to such a model.

While running this study, we were also completing an experiment to test the role of word knowledge in the relationship between different span tasks and verbal abilities. We found that word knowledge is an important factor in these relationships (Engle et al., 1990; see also Dixon, LeFevre, & Twilley, 1988). Controlling for word knowledge by presenting all subjects with very high frequency stimuli did not eliminate the correlations between the span scores and VSAT (although they were reduced). For this study, we do not have an independent assessment of word knowledge, so that this factor may be driving the relationship among some of our variables. We do, however, have two measures for which word knowledge should be completely irrelevant. Neither the STM digit span, nor the probe-recall digit task, involve reading sentences or retaining words, and neither of these relate to the WM measures. By contrast, the STM digit span correlated with the STM word span and both probe-recall tasks, and the probe-recall digit task also correlated with the STM word span. If we assume that a word-knowledge factor drives the correlation between the WM tasks and the simple word tasks, then the rest of the correlation matrix supports the contention that simple memory tasks reflect a different memory system than the complex WM spans. The probe-recall tasks and STM spans interrelate because they both reflect the capacity of STM. Similarly, both WM spans capture the capacity of WM. The correlations between the word versions of the STM measures and both of the WM measures reflect their common reliance on word knowledge, not a common memory system.

Without the assumption that word knowledge has influenced performance on our tasks, it would be difficult to determine from the zero-order correlations, whether our measures reflect the same underlying construct. With a few exceptions, scores on all of the tasks seem to intercorrelate to some extent. Rather than simply speculating about the magnitude of these relationships, we submitted the scores from the tasks represented in Table 3 to a more formal factor analysis. Theoretically, there are several patterns that could emerge from this procedure. For example, all of our tasks may reflect the same underlying limited capacity, but words and digits may simply be processed differently within this system. We might then expect a two-factor solution, one for all of the tasks involving words, and the other for the digit-related tasks. Alternatively, two of our tasks were

presented auditorily and the rest were visual so that we might expect separate factors based on modality differences. It is also possible, in assuming a unitary memory, that we would obtain one factor that represents all of our measures, and one that corresponds only to those tasks that would be influenced by word knowledge. In this case, each task would load on one factor, but the second factor would include only the STM word span, the probe B + C word task, and the working-memory spans. Finally, based on our previous summary of the data, we might expect the analysis is to reveal factors that imply two separate memory systems, one represented by STM and probe-recall measures, and one for the WM spans. The solution of an unweighted least squares analysis with Varimax rotation is presented in Table 4. A two-factor solution was obtained, the first showing its highest loadings from the STM and probe-recall tasks (eigenvalue = 2.62, 43.7% of the variance accounted for), and the second having its highest loadings from the WM measures (eigenvalue = .97, accounting for an additional 16.1%). This outcome clearly supports the notion of separate memories.

Although the factor analysis implies distinct memory systems, there is at least one other reason why a two-factor solution might be found. Procedures for the STM spans and probe-recall tasks are very much alike, each having an experimenter-controlled presentation rate, and neither requiring interpolated activity. Similarly, the WM spans are identical in their procedures. Both are self-paced and require interpolated activity. It is, therefore, possible that what is creating separate factors is variance due to experimental procedures, not different memories. To address this issue, we created composites of each factor, and tested whether they account for different variance in verbal abilities. There is no reason to expect that atheoretical methodological differences would alter the relationship between immediate memory and VSAT. If our measures reflect different memories, then they should capture different variance in VSAT. To test this hypothesis, scores on each of the measures used in the factor analysis were standardized, and tasks loading on each factor were multiplied by their corresponding factor loading and then summed. Omitted was the STM digit span, because it showed no

TABLE 4
Loadings for ULS Factor Analysis With Varimax
Rotation

	Factor 1	Factor 2
STM Word Span	.70	.28
STM Digit Span	.54	.12
Probe Words B + C	.95	.16
Probe Digits B + C	.71	.04
WM Word Span	.19	.82
WM Digit Span	.11	.74

Note. $N = 49$.

relationship to VSAT, and we have argued that it is too influenced by rehearsal to be a good measure of STM capacity. The result was two composite scores for each subject, one for the simple memory tasks (STM word span and both probe-recall tasks), and the other for the two WM measures.¹ A stepwise multiple regression analysis showed that the STM composite significantly predicted VSAT ($R = .54^*$), and that this equation was substantially improved when the working-memory composite was introduced ($R = .65^*$). Although the composites do intercorrelate ($r = .31^*$), part correlations showed that controlling either for the simple memory task composite, or that of the working-memory spans did not eliminate their statistically significant relationships with VSAT ($r = .36^*$, $.40^*$, respectively). This final analysis once again converges on the notion that there are separate memory factors that underlie simple and complex memory spans, and that both of these factors are important to verbal abilities.

DISCUSSION

Prior to summarizing our findings, we wish to emphasize a potentially important point. There are several contemporary models of immediate memory, some focusing on structural aspects of the system (cf. Baddeley, 1986; Baddeley & Hitch, 1974; Monsell, 1985; Schneider & Detweiler, 1988), and others emphasizing its functional properties (Daneman & Carpenter, 1980, 1983; Daneman & Tardiff, 1987). Debate persists as to whether data that implies structural differences actually reflect distinct processes within the same memory (cf., Daneman & Tardiff, 1987; Reisberg, Rappaport, & O'Shaughnessy, 1984). We have, in this article, provided evidence that separate memory tasks do seem to capture separate cognitive "factors," "systems," "mechanisms," or "constructs," but it is not our intent to suggest that these are different structures in the strictest sense. Clearly, our results can generally accommodate models supporting distinctions based on structures or processes, either of which would be interesting.

This study represents part of an ongoing attempt to examine the relationship between STM and WM span tasks, the role of rehearsal in STM tasks, and how all of these factors are manifest in cognitive abilities. We recognize that many of our zero-order correlations are not large, but they are both statistically significant and theoretically consistent. It has been argued that the capacity of STM is important in higher level cognitive tasks, but that simple mnemonics, such as rehearsal, will drive down this relationship (Cohen & Sandberg, 1980; LaPointe & Engle, 1990). In our probe-recall task, we found that performance on beginning-of-the-list items, those that should be most rehearsed, did not correlate with

¹Composites were also computed with weights equal to unity. This produced virtually no change in the results of the subsequent analyses.

VSAT. Middle and final items, which should have received the least opportunity for rehearsal, showed predictive validity. We also expected that digits would be easier to rehearse, so that those STM tasks involving digits should correlate less with VSAT than those involving words. Finding that the STM digit span did not predict VSAT, and that fewer serial positions in the probe-recall digit task correlated with abilities, provides initial evidence for the hypothesis that rehearsal obscures the relationship between STM and complex cognitive abilities. These findings do not, however, speak to the issue of whether or not STM and WM tasks reflect the same underlying construct.

The results of a variety of statistical procedures converged on the conclusion that there are separate memories reflected by our different tasks. Although there is some intercorrelation among all of our tasks that involve words or sentences as stimuli, a factor analysis demonstrated two separate factors, one representing the simple storage tasks and the other defined by the complex WM spans. Moreover, regression and partial correlations showed that each of these factors predicted different variance in VSAT. We proposed that the intercorrelation between some of the STM and WM tasks might reflect the influence of word knowledge, a variable that has been previously shown to be important in memory-span performance (Engle et al., 1990). Interestingly, when the data from Engle et al. were first reported, we assumed that the STM and WM spans reflected the same memory, because they seemed to be influenced by the same variables. We did not, however, directly test this hypothesis. As a result of the outcome in this study, we returned to the data of Engle et al. We reanalyzed that data, selecting the two tasks in which stimulus words were all of very high frequency so that word knowledge should be least influential. Engle et al. reported that both WM and STM spans predicted VSAT ($r = .38^*$, $.45^*$, respectively). In readdressing that data, we have found with partial correlations, that controlling either for the complex mathematical WM span, or the simple STM word span did not eliminate their remaining relationship with VSAT ($r = .32^*$, $.20^*$, respectively, $df = 87$). Again, these are not large correlations, but they are consistent with the data reported in this study, which finds stronger relationships when composites of STM and WM tasks are used, thereby by increasing our predictive power. Taken together, all of our data converge on the conclusion that STM and WM tasks measure different memory mechanisms. These results have important implications for contemporary models of information processing in general, and speak directly to the merits of current frameworks of short-term and working memory.

The most widely cited model of immediate memory is that of Baddeley and Hitch (1974; Baddeley, 1986). The most important component of this framework is the central WM processor, with resources that are required for complex information processing and storage. It is a flexible work space that is not dependent on codes of a particular modality. Verbal STM, often referred to as the "articulatory loop," is more constrained. It is sensitive to word length implying strict temporal parameters (Baddeley, Thompson, & Buchanan, 1975), and the re-

duced STM capacity found with phonologically confusable material suggests that it is particular to verbal coding (Baddeley & Hitch, 1974). Moreover, according to Baddeley and Hitch, STM has its own limited capacity, and if this limitation is not exceeded, then STM can maintain information without requiring central resources.

There is a great deal of data to support the Baddeley and Hitch model (for review see Baddeley, 1986), but it is not without criticism. For example, there is controversy regarding how general the central processor should be viewed, with arguments that there may be separate verbal and nonverbal mechanisms (Daneman, 1987; Daneman & Tardiff, 1987; but see Salthouse et al., 1989). The data reported here cannot distinguish between verbal and nonverbal processors, but within the verbal domain, we have demonstrated evidence for Baddeley and Hitch's model. Simple STM storage tasks, such as the traditional spans and the probe-recall tasks, loaded on a factor that was distinct from complex WM spans. These results imply that there are separate memories (structures or processes). We would argue that the previous studies' finding that the different spans are similarly effected by word knowledge and word length (Engle et al., 1990; LaPointe & Engle, 1990) show that STM and WM may share a sensitivity to certain variables that are important whenever verbal information is processed. With this conclusion, we would expect that there would be some correlation between the two memories even though they reflect different systems. The two composites of our memory tasks did intercorrelate, but also accounted for different variance in verbal abilities, thereby supporting this notion.

Baddeley and his colleagues have focused much of their research efforts on characterizing STM and WM (see Baddeley, 1986, for review) with surprisingly little said about how these memories would manifest in individual differences (but see Cohen & Heath, 1990). There is some research by Baddeley et al. (1985) to suggest that the central processor is a factor in differences in reading ability, but the role of STM in such differences remains vague. Daneman and Carpenter (1980, 1983) on the other hand, provided a model in which memory capacity plays a central role in individual differences. Borrowing from Baddeley and Hitch (1974), they also argued that WM is a computational arena with limited resources for processing and storing information. Although they considered central processing important to all cognitive tasks, they suggested that there will not be a global capacity measure to predict differences in cognitive abilities across a wide range of complex situations. Essentially, Daneman and Carpenter (1980, 1983; Daneman, 1982, 1984) argued that memory spans provide an index of how much information can be stored in WM. But the amount of storage space will depend on how much of the limited resource pool is required for processing. Storage capacity is a by-product of how much capacity is used for processing. Across tasks, each person differs in the efficiency with which information can be processed, so that there are functionally different capacities associated with different tasks. As a result, in order to predict complex abilities from a capacity

measure, the memory task must involve the same processes as those captured by the abilities measure. Finding that the WM reading span predicted a variety of reading-related comprehension measures supported this task-specific view of working memory (Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Masson & Miller, 1983).

Recently, however, the task-specific view of working memory has been challenged by findings that complex WM spans need not be reading-related in order to correlate with verbal comprehension. Performance on complex mathematical spans also show predictive validity (Daneman & Tardiff, 1987; Engle et al., 1990; LaPointe & Engle, 1990; Salthouse et al., 1989; Turner & Engle, 1989). Consequently, Daneman (1987; Daneman & Tardiff, 1987) revised the original task-specific framework, to reflect a "domain-specific" view of immediate memory. She argued that there is a global language-based WM processor that is required for the complex processing and storage of all verbal information. Similarly, there is a nonverbal central processor. As with its predecessor, however, this model does not acknowledge the influence of a dedicated storage system in producing individual differences. According to Daneman (1987), STM spans measure the storage properties of WM when complex processes are not accounted for, and are, therefore, not an appropriate measure of memory capacity as it relates to complex tasks. It has even been argued that simple storage tasks are most interesting in the laboratory, but may have little ecological validity (Daneman, 1987). The data presented here clearly contradict such a hypothesis. It has been demonstrated that individual differences in verbal abilities can be predicted by simple storage tasks, and these tasks are not reflecting the same capacity as that captured by the WM spans. To the extent that verbal abilities are important in everyday life, we would argue that STM is ecologically relevant.

Having argued the merits of distinguishing STM from WM, and concluding that they are both important to individual differences, we now turn to how these factors may function. We generally adopt the framework of Baddeley and Hitch (1974; Baddeley, 1986). WM is viewed as the LTM information active above some critical threshold for current cognitive activity. In terms of Baddeley and Hitch, it is the arena in which sophisticated processing occurs, and where there is on-line storage of information currently being manipulated. STM on the other hand, is a simple storage buffer. It is important during processing, when there is information that does not require immediate manipulation but should be temporarily maintained. Using reading comprehension as an example, creating gists and making inferences about text requires the processing and online updating of information within WM. Remembering the color of a building or the name of a character may also be important, but it is not necessary to use a great deal of attentional resources for this type of maintenance. STM would be better suited for such a task. Consistent with the Baddeley and Hitch (1974) model, STM may be a simple storage buffer whose contents are controlled by a central mechanism, but with a limited capacity for maintaining information without requiring central

attentional resources. We would add to this model that there are important individual differences associated with STM. These are not based on rehearsal or other properties of STM that may be interesting during simple storage situations, but rather reflect the more passive capacity of this system when simple mnemonics or other strategies would require too many attentional resources.

REFERENCES

- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Baddeley, A.D. (1986). *Working memory* (Oxford psychology series No. 11). Oxford: Clarendon Press.
- Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation*. London: Academic.
- Baddeley, A.D., Logie, R., Nimmo-Smith, I., & Breerton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 119-131.
- Baddeley, A.D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 585-589.
- Bower, G.H. (1975). Cognitive psychology: An introduction. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Broadbent, D.E. (1975). The magical number seven after fifteen years. In R.A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory*. New York: Wiley.
- Carpenter, P.A., & Just, M.A. (1988). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*. Hillsdale, NJ: Erlbaum.
- Chiang, A., & Atkinson, R.C. (1976). Individual differences and interrelationships among a select set of cognitive skills. *Memory & Cognition*, 4, 661-672.
- Cohen, R.L., & Heath, M. (1990). The development of serial short-term memory and the articulatory loop hypothesis. *Intelligence*, 14, 151-171.
- Cohen, R.L., & Sandberg, T. (1977). Relation between intelligence and short-term memory. *Cognitive Psychology*, 9, 534-554.
- Cohen, R.L., & Sandberg, T. (1980). Intelligence and short-term memory: A clandestine relationship. *Intelligence*, 4, 319-331.
- Daneman, M. (1982). The measurement of reading comprehension: How not to trade construct validity for predictive power. *Intelligence*, 6, 331-345.
- Daneman, M. (1984). Why some people are better readers than others: A process and storage account. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2). Hillsdale, NJ: Erlbaum.
- Daneman, M. (1987). Reading and working memory. In J.R. Beech & A.M. Colley (Eds.), *Cognitive approaches to reading*. New York: Wiley.
- Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Daneman, M., & Carpenter, P. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 561-583.
- Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25, 1-18.
- Daneman, M., & Tardiff, T. (1987). Working memory and reading skill reexamined. In M. Coltheart (Ed.), *Attention and Performance XII*. London: Erlbaum.
- Das, J.P., & Sui, I. (1989). Good and poor readers' word-naming time, memory span, and story recall. *Journal of Experimental Education*, 52, 101-114.

- Dempster, F. (1981). Memory spans: Sources of individual and developmental differences. *Psychological Bulletin*, 89, 63-100.
- Dempster, F., & Cooney, J.B. (1982). Individual differences in digit span, susceptibility to proactive interference, and aptitude/achievement test scores. *Intelligence*, 6, 399-416.
- Dixon, P., LeFevre, J., & Twilley, L.C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, 80, 465-472.
- Engle, R.W., Nations, J., & Cantor, J. (1990). Is "working-memory capacity" just another name for word knowledge? *Journal of Educational Psychology*, 82, 799-804.
- Francis, W.N., & Kucera, H. (1982). *Frequency analysis of English language*. Boston, MA: Houghton Mifflin.
- 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.
- Larson, G.E., & Alderton, D.L. (1988). *Cognitive frameworks for g*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Larson, G.E., & Alderton, D.L. (1990). The structure and capacity of thought: Some comments on the cognitive underpinnings of *g*. In D.K. Detterman (Ed.), *Current topics in human intelligence: Vol. 2. Is mind modular or unitary?* Norwood, NJ: Ablex.
- Masson, M.E.J., & Miller, J.A. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology*, 75, 314-318.
- Merkel, S.P., & Hall, C. (1982). The relationship between memory for order and other cognitive tasks. *Intelligence*, 6, 427-441.
- Monsell, S. (1985). Components of working memory underlying verbal skill: A "distributed capacities" view. In H. Bouma & D.G. Bouswhuis (Eds.), *Attention and Performance X*. Hillsdale, NJ: Erlbaum.
- Perfetti, C.A., & Lesgold, A. (1977). Discourse comprehension and sources of individual differences. In M.A. Just & P.A. Carpenter (Eds.), *Cognitive processes in comprehension*. New York: Wiley.
- Reisberg, D., Rappaport, I., & O'Shaughnessy, M. (1984). Limits of working memory: The digit-digit span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 203-221.
- Salthouse, T.A., Mitchell, D.R.D., Skovronek, E., & Babcock, R.L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 507-516.
- Schneider, W., & Detweiler, M. (1988). A connectionist/control architecture for working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21). New York: Academic.
- Turner, M.L., & Engle, R.W. (1989). Is working-memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.