WORKING MEMORY CAPACITY

Marilyn L. Turner and Randall W. Engle
University of South Carolina
Columbia, S. C.

ABSTRACT

Recent researchers have attempted to correlate measures of working memory (WM) with measures of higher level cognitive skills and abilities focusing on the functions of this limited capacity system, i.e., processing and storage. Relationships between three span measures of the functional model of WM capacity and two measures of reading comprehension were investigated. The magnitude of the correlations found between reading comprehension and the two spans embedded in reading processing tasks was similar to that of the correlation found between a third span measure embedded in a quantitative task with reading comprehension. These results indicated that these span measures of WM capacity were independent of the nature of the concurrent processing task.

Higher level cognitive functioning makes demands on a limited capacity system and most memory theorists assign this function to a short term memory (STM) structure. Baddeley & Hitch (1974) pointed out that attempts to relate traditional STM measures with measures of higher level functioning, such as reasoning, learning and reading comprehension have not been successful and that any theory proposing a central role for STM in higher level functioning must focus on a functional analysis, i.e., a working memory (WM). They proposed a model of WM that includes both structural and processing components. The three structural components are: (1) a central executive, (2) an articulatory loop, and (3) a visuo-spatial scratch pad. The articulatory loop and visuo-spatial scratch pad are considered maintenance systems controlled by the central executive. They view the central executive as a flexible work space with limited capacity, with part of the capacity used for processing incoming information and the remainder used for storage of the products of that processing. Baddeley and his collegues presented evidence for the existence of their proposed WM slave components; the articulatory rehearsal loop (Baddeley & Hitch, 1974; Baddeley, 1979; Salame & Baddeley, 1982), and the visuo-spatial scratch pad (Baddeley & Lieberman, 1980).

Although other models of WM have been developed (e.g., Klapp, Marshburn & Lester, 1983), virtually every conceptualization of WM assumes that there is a limitation in the amount of information that can be kept active at any given time. Further, it is generally assumed that this limitation affects consequent processing, i.e., that higher level processing is limited to some extent by the limitations of WM. Beyond these generalizations about the WM there is disagreement about the nature of the WM capacity limitation. Our pilot study begins an investigation of WM, specifically seeking to answer the following question regarding the nature of WM capacity: Is WM capacity invariant within the individual or does it vary with the task as a function of the individual's skill in the task being performed at the time?

Daneman & Carpenter (1980) investigated the role WM plays in reading comprehension. They suggested that, while reading, an individual activates specific strategies and skills peculiar to the reading task. While activated, these strategies occupy WM and are used when processing incoming information in WM, which leads to intermediate products which must also be stored in WM. Daneman & Carpenter hypothesized that individual differences in reading comprehension could be due to readers having different levels of processing efficiency. They developed a complex reading span measure that was a combination of two concurrent tasks, requiring processing and storage functions. In a secondary reading task subjects read aloud series of unrelated sentences which limited their resources for the concurrent critical memory task, i.e., sequential recall of the last word of each sentence in the series. The number of words readers recalled against the background of this reading task was considered their reading span, a measure of their WM storage capacity. The reading span correlated significantly with three measures of reading comprehension, including individuals' Verbal SAT scores, while a traditional word span measure did not. In several studies Daneman & Carpenter (1980, 1983) found high correlations between reading span and Verbal SAT ranging from .49 to .59, and also between reading span and their measure of reading comprehension from .72 to .90. Their interpretation of these findings suggests that good readers have large WM capacities because they have efficient reading strategies, and therefore more efficient processing, leaving more of the total WM capacity for the storage of memory items. However, it should be noted that this suggestion would assume that the secondary concurrent task, i.e., reading unrelated sentences, used in the predicting span measure must be highly related to the criterion measure, i.e., reading comprehension. Reading spans are assumed to reflect differences in residual WM capacity because of differences in the efficiency of reading skills, NOT because of differences in total WM capacity independent of task proficiency. Another important possibility, however, is that one of the reasons people are good readers is because they have a larger total WM capacity. According to this view, a WM index should successfully transcend secondary task dependency in its prediction of reading comprehension, or any higher level cognitive skill. While investigating individual differences in adult reading comprehension, Baddeley, Logie, Nimmo-Smith & Brereton (1985) were also interested in the interpretation of the significant correlation between comprehension and the reading span measure. They attempted to test how general the WM system is that underlies the span measure by comparing its correlation with reading comprehension with that of a counting span measure and comprehension. Although, the counting span did not have the same strong relationship with reading comprehension as the reading span did, Baddeley, et al (1985) pointed out that other WM investigations may find a better measure of general WM capacity that is not task dependent.

The present study is directed at testing whether a strong relationship between WM span measures and reading comprehension can occur with span measures requiring different strategies and skills than those required in reading comprehension. Using a task involving strings of arithmetic operations followed by a to-be-remembered word, would allow the memory span test to be embedded in a concurrent task that is seemingly not highly related to reading skills, yet would induce considerable processing demands concurrent with the memory task. Hitch (1978) found that solving mental arithmetic problems utilized processing strategies, which included a temporary storage of initial and interim information. However, the processing strategies used for reading sentences are inherently different than those used in solving arithmetic operations. If the magnitude of the correlation between operation span and reading comprehension is similar to that between reading span and comprehension, then this complex span index may be tapping something more basic than a specific skill-level based system. At the least, it would suggest that individual differences exist in WM capacity independent of specific strategies used in the processing task in which the memory span test is embedded. The stronger implications would be that individual differences in these measures of WM may indicate differences in total WM capacity independent of the task being performed.

This study also seeks to control for a potential confounding in Daneman & Carpenter's (1980, 1983) experiments. Good readers may have remembered more words in the reading span task simply because they comprehended the sentences better than poor readers. At the time of recall, better retention of the gist of the sentences by the good readers could be used to reconstruct the final sentence words. If so, any correlation found between the reading span and comprehension measures would simply mean that the ability to comprehend correlates with the ability to comprehend. A reading span task wherein unrelated sentences are followed by a to-be-remembered digit could correct for this problem, i.e., a sentence digit task. If the relationship between the sentence digit span and reading comprehension still holds, then one could consider (as Daneman & Carpenter (1980) did) that these memory span measures are indexing a limited capacity WM.

In summary, this study addresses two issues by testing subjects with three span tasks. First, a sentence-word task is used which replicates the task used by Daneman & Carpenter (1980) in which subjects read unrelated sentences wherein the span measure is the number of last words recalled (SW). In addition, a sentence-digit task is used in which subjects read sentences followed by to-be-remembered digits (SD). Comparing the correlation of the spans resulting from the two tasks (i.e., SW & SD) with reading comprehension measures should clarify whether Daneman & Carpenter's (1980, 1983) reading span (SW) is a measure of what is actually remembered, or is a measure of reading comprehension in disguise. More importantly, an operation-word span task is used in which subjects perform arithmetic operations followed by a to-be-remembered word (OW). Comparing the correlations of the OW span and the two reading spans, i.e., SW and SD, with measures of reading comprehension and Verbal SAT scores should help determine whether these embedded memory spans are reflecting individual differences in the capacity of WM independent of processing task, or are reflecting individual differences in task-specific processing.

METHOD

Subjects

Thirty-seven students from the University of South Carolina participated, fulfilling a requirement for a psychology course. Each subject was seen individually, and over a period of approximately one hour participated in a reading comprehension task and three complex memory span tasks. Also, Verbal SAT scores were obtained with subjects' permission.

Reading Comprehension Task

Materials. Twelve paragraphs, each of which was approximately 160 words in length, were accompanied by four multiple choice or short answer questions.

Procedure. Subjects were instructed to silently read the paragraphs and to be prepared to answer factual questions about each paragraph immediately after reading it. No time limits were imposed for answering the questions or for reading the paragraphs, and subjects were allowed to leave an answer blank if they could not recall the answer.

Memory Span Tasks

Materials. The set of materials for the sentence-word (SW) task consisted of 60 unrelated sentences presented on flash cards to the subject. There was one sentence from 11 to 14 words long on each card. The number of cards per trial varied from 2 to 6. Blank cards separated the trials and cued subjects' serial recall. The complete card deck consisted of a specific sequence of 15 trials which gradually increased the number of sentences (cards) from two to six in each sequence of trials. There were three trials for each of the sequentially increasing series (2-6) thus, there was a total of 60 sentences used in all fifteen trials. The set of cards for the sentence-digit (SD) task was identical to the SW set with the exception that the sentences in this set were followed by one-digit numbers, randomly sampled with replacement from numbers 0-9. Each card in the set for the operation-word (OW) task consisted of an unanswered numerical operation string (e.g., (3 ■ 4) + 11 = ...) followed by a randomly chosen word (e.g., apron, received, etc.). The first arithmetic problem in all operation strings consisted of one single-digit multiplication or division operation in parentheses, such as (7 ■ 1) or (8 / 2), with each digit randomly sampled with replacement from the numbers 0-9. This problem always resulted in a whole number and was to be solved prior to the following addition or subtraction problem, consisting of a one- or two-digit number, such as, + 13, - 3 or - 19, randomly sampled with replacement from the numbers 1-19. The length of each word following an operation string varied from one to three syllables and was randomly chosen from English literature text books for college freshman level courses without regard to any other criteria.

Procedure. Each subject was instructed to read aloud all the information on each card, including the target item. For example, in the SW and SD tasks subjects read the sentence on each card aloud, and when a blank card occurred serially recalled the last word (or digit) following each of the sentences in that trial. In the OW task subjects read the operation aloud, then solved the operation giving the answer aloud, and finally read the target word aloud. Audible and/or silent intermediate computations were permitted, but no pencil and paper aids. For all memory span tasks when the subject successfully recalled the target items in two of the three trials for a specific size series, the size of the series was incremented. Subjects were warned to expect the number of cards to increase as progress was made from a 2-card to a 6-card series, and subjects proceeded through the series of each stimulus set until they failed two of the three trials in a series. A separate memory span for each stimulus set was defined at the highest series size wherein a subject recalled at least two of the three trials correctly.

RESULTS

Table 1 shows Pearson Product Moment correlations found between each of the three span measures with reading comprehension and Verbal SAT scores.

TABLE 1

SPAN MEASURES	COMPREHENSION MEASURES	
	Paragraph Questions	Verbal SAT
Sentence-word	.36	.39
Sentence-digit	.49	.50
Operation-word	.41	.38

N=37 for all three span measures.
N=37 for paragraph question measure.
N=35 for Verbal SAT measure.

The similar correlations found between the two comprehension measures and the three span measures embedded in arithmetic and reading related tasks clearly indicated that the spans are independent of specific skills required in the concurrent processing task. It can be seen that the highest span-comprehension correlations were between sentence-digit (SD) span and Verbal SAT scores ($r(33) = .50$, $p<.01$), and between SD span and our reading comprehension measure ($r(35) = .49$, $p<.01$). Table 1 also shows that sentence-word (SW) and operation-word (OW) spans correlated significantly and nearly equally with the two comprehension measures. Subjects with higher SW and OW spans were better at answering paragraph questions than subjects with lower SW and OW spans, $r(35) = .36$ and $.41$, $p<.02$ and $.01$ respectively, and, subjects with higher SW and OW spans achieved higher Verbal Sat scores, $r(33) = .39$ and $.38$, $p < .02$ respectively. These results indicate that it is NOT necessary for the predictor measures, i.e., the complex memory spans, to be embedded in concurrent processing tasks requiring skills related to the criterion measure, in order to find significant correlations between them.

The sample distribution of Verbal SAT scores ranged from 250 to 640 (mean = 459, sd = 96), and the distribution of our second measure of comprehension, i.e., answers to the paragraph questions, ranged from 37.5% to 91.67% with a mean of 70.16% (sd = 13.23%).

Table 2 shows the frequency distributions of the three span measures. While Daneman & Carpenter (1980) did not present a frequency distribution of their span scores, the information they did provide suggests our sample is distributed somewhat differently than theirs.

TABLE 2

SPAN MEASURE	SPAN SCORE				
	2	3	4	5	6
SENTENCE-WORD	17	16	3	0	1
OPERATION-WORD	15	18	1	0	3
SENTENCE-DIGIT	1	4	9	12	11

Specifically, the SW span scores, resulting from our replication of Daneman & Carpenter's reading span task, revealed a positive skewness in an already limited range. With the exception of one subject, the range of our SW spans was between 2 - 4 (mean = 2.70, sd. = .85), while Daneman & Carpenter's spans ranged between 2 and 6, (mean = 3.15, sd. = .93). Since our subjects had the more limited range in the reading span measure, comparison of our results with theirs should be modified by the possibility that our large span subjects may be more like their "better" small span subjects.

Subjects were better at remembering digits than words. The sentence-digit (SD) spans averaged 4.76 (sd = 1.09), while the sentence-word spans averaged 2.70 (sd = .85) and the operation-word (OW) spans averaged 2.89 (sd = 1.07). While both the SW and OW spans revealed positively skewed distributions, the SD span distribution was negatively skewed. Thus, our span distributions are plagued with a restricted range and extreme skewness and would seem to bias against finding the correlations between the measures we designed the study to test, and measures of reading comprehension.

A secondary issue addressed by this paper was whether Daneman & Carpenter's reading span is simply a measure of reading comprehension in disguise. A comparison of the correlations between SW span and comprehension with those between SD span and comprehension answers this question. Both of these span measures involved subjects reading unrelated sentences, but the SW task required recall of last words of the sentences, while the SD task required recall SD task required recall of digits following the sentences. Table 1 shows that SW spans correlated .36 with answering paragraph questions and .39 with Verbal SAT scores, and that SD spans correlated .49 with answering paragraph questions and .50 with Verbal SAT scores. Comparing the correlations between the paragraph questions and the SW measure ($r= .36$) and the SD measure ($r= .49$) showed the two correlations were not significantly different, $t(36)$

= .84, p > .20. Also, when comparing the correlations between these two span measures and Verbal SAT scores (r = .39 and .50 respectively) no significant difference was found, t(36) = .61, p > .20.

DISCUSSION

The basic question addressed by this study was whether individual differences arise from the use of more or less efficient processing strategies, or, from an inherent difference in WM capacity. Investigating this issue, we asked if the predictor WM span was measured using a processing task which required the same specific reading related skills that are used in the criterion measure, i.e., reading comprehension. The similar correlations found between the reading comprehension measures and the span measures embedded in the arithmetic and reading processing tasks clearly indicated that the measures were independent of the type of processing task. All three span measures reliably predicted comprehension, independent of any specific skills required in any particular processing task performed concurrently with the span task.

In addition, a comparison of the correlations between SW and SD span measures with the reading comprehension measures argues against the notion that the reading span measure used by Daneman & Carpenter, (i.e., our SW span) is contaminated by the subject simply recalling the gist of the individual sentences and reconstructing the final words from that. Although reconstruction of the gist of the sentences could have facilitated recall of the last words of each sentence in the SW span task, this gist reconstruction could not possibly have been used to reconstruct the digits that followed each unrelated sentence in the SD span task. The concurrent processing task of reading a series of unrelated sentences while holding the last word of each sentence in memory for future recall is not a measure of reading comprehension in disguise.

The current finding that the embedded span measure is independent of task, i.e., that the sentence-word (SW), the sentence-digit (SD) and the operation-word (OW) span measures all showed similar correlations with both measures of reading comprehension, clearly shows that the span measure does not depend on specific processing strategies. Although, we wish to caution the reader that these data are correlational in nature, and therefore cause-and-effect attributions cannot be made, we can conclude that the embedded memory spans are reflecting individual differences in the capacity of working memory independent of processing task, strongly implying that differential reading strategies and reading skills are not the only factor determining the WM span measure. There must be a more general system underlying the WM span measure.

REFERENCES

Baddeley, A.D. (1979). Working memory and reading. in P.A. Kolers, M.E. Wrolstad, & H. Bouma (Eds.), Processing of Visible Language. 355-370, NY: Plenum.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In G. H. Bower (Eds.), The Psychology of Learning and Motivation, 8, New York: Academic Press.

Baddeley, A.D. & Lieberman, K. (1980). Spatial working memory. In R. Nickerson (Ed.), Attention and performance, 8, 521-539, NJ: Lawrence Erlbaum Associates.

Baddeley, A.D., Logie, R., Nimmo-Smith, I. & Brereton, N. (1985). Components of fluent reading. Journal of Memory and Language, 24, 119-131.

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

Daneman, M. & Carpenter, P.A. (1983). Individual differences in integrating information between and within sentences. Journal of Experimental Psychology: Learning Memory & Cognition, 9, 561-584.

Hitch, G.J. (1978). The role of short-term working memory in mental arithmetic. Cognitive Psychology, 10, 302-323.

Klapp, S.T., Marshburn, E.A., & Lester, P.T. (1983). Short term memory does not involve the "working memory" of intellectual processing: the demise of a common assumption. Journal of Experimental Psychology: General, 112, 240-264.

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