What Do Estimates of Working Memory Capacity Tell Us?

Nelson Cowan, Candice C. Morey, Zhijian Chen, and Michael F. Bunting

Department of Psychological Sciences

University of Missouri

18 McAlester Hall

Columbia, MO 65211 USA

E-mail: CowanN@missouri.edu

Telephone: 573-882-4232

Abstract

We sketch our recent research leading toward a revision of the concept of working memory capacity. This includes a rationale for believing that working memory capacity is a fundamental ability, and for estimating working memory based on the number of independent chunks of information recalled. We provide evidence that the often-noted relation between working memory and both individual and developmental differences in intellectual aptitudes does not depend on using a dual task to measure working memory, even though that has become a standard practice in the field. Basic intellectual aptitudes can be studied more cleanly using working-memory tasks that impede mnemonic strategies, so that a core capacity must come into play. Last, we offer evidence that this core capacity may be attention serving as a storage device.

What Do Estimates of Working Memory Capacity Tell Us?

Working memory can be viewed as the collection of mental processes that preserve a limited amount of information in an especially accessible form, long enough for it to be of use in ongoing cognitive tasks. By almost all accounts, working memory is indeed a collection of processes (e.g., for diverse theoretical descriptions of working memory, see the chapters of Miyake & Shah 1999). In our laboratory, we have been investigating "flavors" of working memory that we believe to be closely related to human conscious awareness, namely the *primary memory* described as the trailing edge of consciousness by James (1890). This chapter serves as a sort of tour of the laboratory's current endeavors, made more or less interactive by the inclusion of a considerable amount of data from recently-published studies, presented here in tabular form so it can be explored and questioned

We examine research that makes four related points in progression: (1) that working memory is an important concept that amounts to more than some notion of general processing efficiency; (2) that the often-discussed relation between working memory and intellectual maturation and aptitude does not depend on using a dual task to examine working memory, but instead depends only on using a task that impedes mnemonic strategies such as covert verbal rehearsal; (3) that working memory performance depends on the notion of capacity, expressed in terms of chunks; and (4) that the core working-memory capacity limit is related to the scope of attention.

Premise 1: Working Memory Is Not Just General Processing Efficiency

There is a tradition in the research literature suggesting that developmental differences in working memory and cognitive aptitudes throughout the life span result from a developmental

increase, and then decrease, in the speed of processing as neural efficiency improves with maturation in childhood and later declines with aging (e.g., Kail & Salthouse 1994). The theoretical hypothesis is that individuals with a better working memory are able to reactivate information more quickly before it can decay from working memory, and therefore that the speed of processing may be the key factor distinguishing between people with better or poorer working memory. The basis of this claim was a wealth of evidence that processing speed and the accuracy of performance are highly correlated across individuals.

Cowan, Wood, Wood, Keller, Nugent, and Keller (1998) documented a strong correlation between two measures of processing efficiency and performance on a digit span task. The digit span task is one of the oldest measures of the short-term retention processes that form a key component of working memory, and is included in standard tests of intelligence. In this task, a series of digits between 1 and 9, with each digit used no more than once per list, is presented on each trial. The list length starts very short and increases every few trials, with the exact number of trials per list length differing from one procedure to another. If the participant is incorrect on every trial at a given length, the test stops, and therefore the duration of the test is adjusted to the abilities of the participant. Some measure of the length of list that can be repeated is taken as an individual's digit span.

One measure of efficiency examined by Cowan et al. (1998) was the rate at which digits were pronounced in lists of a particular length within the span task. The durations of every word in the spoken responses, and the durations of silent pauses between words in these responses, were painstakingly measured using a computer. An oscillographic display of the response was measured with auditory guidance. The notion was that individuals who process information faster would repeat more of this information before it decays. Another measure of efficiency

was one devised by Baddeley, Thomson, and Buchanan (1975). In this type of measure, a small number of items is to be repeated aloud as quickly as possible, either with a set short enough to be memorized and repeated over and over or with a set that is read rather than remembered. This type of measure was thought to estimate how quickly the participant is capable of covertly rehearsing a longer list of words drawn from the same set, in a span task.

Cowan et al. (1998) found that both the speed of retrieval and the speed of rehearsal were significantly related to span. A statistical method (structural equation modeling) was used to characterize the relations between measures. It was found, contrary to the expectations of a single type of efficiency across tasks, that retrieval and rehearsal speed were unrelated to each other. However, they both were related to digit span, to an extent described in Table 1. These were both strong, separate relations and together they accounted for 87% of the age-related variance in digit span (and 60% of its total variance).

- - - - - Table 1 - - - - -

At the time, this finding provided one of the strongest arguments in favor of some sort of efficiency theory of working memory, although it required that one stipulate the presence of more than one efficiency parameter per individual (because retrieval and rehearsal rates did not correlate with one another). However, it always bears repeating that correlations do not amount to evidence of causation, even with the use of structural equation modeling.

Recently, Cowan et al. (in press) showed that at least the retrieval rate factor is not causally related to digit span. Ordinarily, children who are 8 to 9 years old have a smaller digit span than adults and retrieve items (pronounce them in the response) much more slowly. In two different ways within separate experiments, Cowan et al. were able to get the children in an experimental group to recall items more quickly. In Experiment 1, this was accomplished by speeding up the

stimuli. In Experiment 2, one group of children was simply instructed to speak more quickly in their responses at a certain point (Phase 2 in Table 2), and the children were able to do so. In fact, these children sped up their responses to be faster than adults usually speak when recalling lists of an equivalent length. Yet, as illustrated in Table 2, speeding up did not affect span at all. The quickly-speaking children still were just as far behind the adults in digit span. This study emphasizes that the correlation between efficiency or speed and span does not reflect a forward direction of causation. Perhaps, instead, the most comfortable speed of responding and memory capacity are separate factors that increase during childhood. By analogy, the head and the arms both grow larger during childhood, but one instance of growth does not cause the other; stretching out the arms as far as they can go does not affect the size of the head.

- - - - - Table 2 - - - - -

Premise 2: The Relation Between Working Memory and Intellectual Maturation and Aptitude Does Not Depend on Using a Dual Task to Examine Working Memory

The traditional view of working memory (Baddeley & Hitch 1974; Baddeley 1986; Baddeley & Logie 1999) and its development (Gathercole & Hitch 1993; Hitch, Towse, & Hutton 2001) is one that incorporates storage and processing mechanisms that are assumed to be distinct , both neurologically and behaviorally. From that viewpoint, it stands to reason that, to test the capabilities of working memory, the test should tax or engage both storage and processing. That has been the logical underlying the type of test that has become prevalent in the literature to examine individual and developmental differences in working memory. In the *reading and listening span* tests (Daneman & Carpenter 1980), a sentence is presented and some comprehension is required. The participant is also to remember the last word in the sentence or, in a variation of this procedure, is to remember a separate word. After two or more such

sentences, all of the final words are to be repeated. In the *counting span* test (Case, Kurland, & Goldberg 1982), the participant is to count arrays of dots while remembering the sums, and then to repeat all of the sums at the end of the trial. In the *operation span* test (Turner & Engle 1989), the participant is to complete multiple arithmetic problems, remembering either the answers to the problems or a separate word presented after each problem. The number of sentences, arrays, or problems that can be processed along with successful serial repetition of the items following the processes (indicating that these items were successfully stored) is taken as the working memory span. In sum, the structure of a trial in these studies is:

Process 1, Store Item 1, Process 2, Store Item 2....Process n, Store Item n; Recall of Items

The finding has been that this storage-and-processing type of test generally yields correlations with aptitudes that are considerably stronger and more consistent than the correlations between simple span tasks and aptitudes, and that are general across the processing domains (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle in press; Daneman & Merikle 1996; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle 2004). At times, the correlations have been strong enough to suggest that working memory may be the key factor distinguishing between individuals with higher or lower intelligence (Conway, Cowan, Bunting, Therriault, & Minkoff 2002; Engle, Tuholski, Laughlin, & Conway 1999; Kyllonen & Christal 1990).

Nevertheless, there are problems with this approach. These correlations are not useful in understanding intelligence unless we can understand what mechanisms are indexed by the tests of working memory. Perhaps the high correlations occur only because the working-memory tasks require skills other than the temporary storage of information. For example, consider the original version of the listening and reading span tasks, in which the last word of every sentence is to be remembered (Daneman & Carpenter 1980). It seems theoretically possible that this can

be carried out by using long-term memory to retrieve the sentences, using their meaning or grammatical structure as cues. These remembered sentences would serve in turn as strong cues to the last word in each sentence.

This hypothesized retrieval method would lead to the prediction that response times should be much longer for listening or reading spans than for other types of working memory test, in which there are no strong linguistic cues. Indeed, this was found to be the case, by Cowan et al. (2003). For example, Table 3 shows the mean response durations to two-item lists for three types of span in three age groups. Clearly, in all age groups, responses were longer in listening span than in digit span or counting span, and this difference between tasks was exaggerated in young children. Based on these findings alone, one might consider these original versions of listening and reading span to be potentially invalid measures of working memory, because a long-term retrieval strategy is involved. However, it does not necessarily invalidate the other measures of working memory that incorporate separate processing and storage components.

- - - - - Table 3 - - - - -

Another possible drawback of storage-and-processing tasks is that they require dual-task coordination. For some individuals at least, this poses a difficulty that depends on central executive functioning and can be separate from working-memory storage (Logie, Cocchini, Della Sala, & Baddeley 2004). Fortunately, though, the incorporation of separate storage and processing components may not be necessary after all. Recent studies lead to a reinterpretation of why they correlate so well with aptitudes; that is, why they "work."

The original interpretation was that both processing and storage must be involved in a working memory task for it to work well. An alternative interpretation is that the processing episodes merely prevent covert rehearsal, which can become relatively automatic in adults

(Guttentag 1984) and can circumvent the need to expend a basic working memory capacity on storage per se. Thus, the difficulty of the processing component of the working memory task is not particularly critical (Duff & Logie 2001). The effectiveness of the working memory task instead seems to depend on how tight the schedule of retrieval and central processing is (Barrouillet, Bernardin, & Camos 2004; Conlin, Gathercole, & Adams 2005; Friedman & Miyake 2004; Lépine, Barrouillet, & Camos 2005), so that it can leave little time for rehearsals to sneak in between processing episodes.

As emphasized by Cowan (2001), there are some sorts of working memory task in which rehearsal is not feasible despite the absence of a dual task. In some of these tasks, an array of items is presented too quickly for rehearsal to be possible. A classic example is the array memory procedure of Sperling (1960), in which a briefly-presented array of characters was followed by a partial-report cue. A more recent example is the visual array comparison task of Luck and Vogel (1997), in which an array of small, differently-colored squares is followed by a second array, within about 1 sec of the first, that matches the first array or differs in the color of one square. In one version of the task, to make the decision easier, the second array includes one encircled square and, if any square changed color, it was that one. The task is to indicate whether the square changed color. This task is unaffected by the exact duration of the brief, first array and unaffected by the need to recite digits during the task to suppress, provided that the recited digits do not impose a substantial memory load (Luck & Vogel 1997; Morey & Cowan 2004, 2005). Adults can carry out this task almost perfectly with 3 or 4 squares in each array, but performance drops dramatically as the number of squares increases beyond working-memory capacity.

Another example of a working memory task that does not include a separate processing

component, but still makes rehearsal infeasible under some circumstances, is running memory span (Pollack, Johnson, & Knaff 1959). In this sort of task, a long list of items ends unpredictably, after which the task is to recall as many items as possible (or a certain requested number of items) from the end of the list. If the list is presented very quickly (3 or 4 words per sec, which can be done with spoken words without losing intelligibility), it is impossible to rehearse. In fact, Hockey (1973) found that, with rapid presentation in a running span task, rehearsal instructions led to *worse* performance than instructions to wait passively for the list to end. (With a slow presentation, in contrast, rehearsal instructions help.)

Cowan, Elliott et al. (2005) recently showed that working memory tasks that do not allow much rehearsal, but nevertheless do not include separate storage and processing components, do very well in accounting for variance in aptitude tests. They did not do quite as well as listening span or counting span tests, but the extra variance accounted for by those tests appeared to be mostly task-specific and not general across both types of storage-and-processing task. Therefore, the tasks such as visual array comparisons and running span could be viewed as purer indicants of working memory capacity. It, like the storage-and-processing tasks, did much better than simple digit span in accounting for aptitude test results.

An additional expectation could be drawn from our analysis of this single-component type of working-memory task. Young children are typically unable to rehearse automatically the way that adults do. Therefore, in young children, even simple digit spans should show a high correlation with aptitudes. This is exactly what Cowan, Elliott et al. (2005) found, as is shown in Table 4. In second- and fourth-grade children, digit span had a high value in predicting a general factor of intelligence extracted across verbal and nonverbal tests. In sixth-grade children and adults, in contrast, digit span was of no predictive value.

- - - - - Table 4 - - - - -

Premise 3: Working Memory Performance Depends on the Notion of Capacity, Expressed in

Terms of Chunks

This discussion of what types of test predict aptitudes leads to the question of what core

capacity exists, and plays a role in aptitude. In a famous paper, Miller (1956) noted that people

seem able to recall about seven items in a span test, plus or minus a couple. However, in the

same paper he emphasized the importance of the ability to group items together to form higher-

level, meaningful units or *chunks*. As a compelling example (not used by Miller), the letter

string BBCCIAFBI would be difficult to remember as nine separate letters but it is rather easy to

remember if one recognizes three acronyms for well-known agencies, the British Broadcasting

Corporation (BBC), the Central Intelligence Agency (CIA), and the Federal Bureau of

Investigation (FBI). Then it remains possible that the reason people can remember about seven

items is that they rapidly form new, larger chunks of information. The reason why telephone

numbers are typically presented in groups of three and four digits is probably to assist in the

formation of new chunks. It is also possible that rehearsal of the digits in a repeating loop

(Baddeley 1986) is used to form these chunks.

One of the founding fathers of the field of cognitive psychology, Donald Broadbent,

recognized that there is probably a core capacity that remains after mnemonic strategies like

grouping and rehearsal do not operate (Broadbent 1975). He suggested that this core capacity is

three items. (Conceptually, three chunks given an absence of grouping.) He gave the example

of the list length resulting in perfect performance; usually, across many trials, no more than three

items. Another example he gave was that recall from long-term memory requires repeatedly re-

filling and emptying working memory, as a conduit between the long-term store and voluntary

response processes.   For example, when one is asked to recall all of the states of the U.S., the answers come out in bursts of just a few states at a time.

Cowan (2001) more systematically looked for situations in which each item remains a separate chunk because of aspects of the task that prevent grouping and rehearsal. Some of these situations are summarized in Table 5. What was remarkable from this survey was that diverse types of situation converged on an estimate of 3 to 5 items recalled in these circumstances.

- - - - - Table 5 - - - - -

There is something less than completely satisfying about characterizing the number of chunks that can be recalled entirely on the basis of situations in which each item equals a chunk. How do we know that the task analyses are correct, and how do we know whether the same estimate of capacity (3 to 5 chunks) would emerge in situations in which multiple items are grouped together to form a chunk? We have investigated this question intensively in our recent work. There already was some evidence that people can remember 3 to 5 chunks in a situation in which they are to replicate a chessboard, with pauses between groups of pieces defining a chunk (Gobet & Simon 1998). We hoped to extend the evidence to verbal recall, in which there is strong evidence of chunking processes (e.g., Johnson 1978; Marmurek & Johnson 1978).

Tulving and Patkau (1962) already found that presenting material with associations between items resulted in the recall of larger chunks, but not more chunks. Stimulus lists were series of 24 words that varied in the order of approximation to English, ranging from random words at one extreme to perfect English sentences at the other extreme. In between, the orders of approximation made sense for any n words in a row, where n increased with the order of approximation, but did not make sense for the entire sentence. For example, in a third-order approximation, each sequence of three words is valid, as in *Today I went home with them out or*

in our house is cold in winter we think.... Free recall was required and the results were to be scored according to serial order: Any series of words recalled in the presented order was considered a single chunk. So if the phrase *Think before you speak up again* were recalled as *up again, you speak,* it would be scored as the recall of two chunks. It was found that the average chunk grew larger with increasing levels of approximation to English but that, at all levels of approximation, 4 to 6 chunks were recalled. There are, though, several shortcomings of this method. Each series may not really be a single chunk in the participant's representation, underestimating the number of chunks recalled; and, conversely, an overall representation of some of the ideas might be used, especially in the higher-order approximations to English, reducing the number of words or chunks that had to be separately retained in working memory. Still, the results show at least a rough agreement with estimates of limited capacity that fall between three items (Broadbent 1975) and seven items (Miller 1956).

Cowan, Chen, and Rouder (2004) tried to improve upon this sort of method. Monosyllabic words were exposed to participants sometimes as singletons and sometimes in pairs, to manipulate the strength of the learned pairing. For example, the words *brick* and *hat* might be presented sometimes in isolation and other times, for the same participants, as the consistent pair *brick-hat*. After a large number of presentations of this sort, 8-word lists were presented, each of which had been constructed from pairs that had received the same strength of pairing, in a way that maintained the learned pairings. There were five list types, shown in Table 6. Recall was scored in terms of whether a pair of words was retrieved in the order that had been presented in the list (an intact pair). Items recalled not in an intact pair were considered singletons; these comprised words recalled without the paired word, or word pairs recalled with its two items in the reversed order and/or separated by other words.

- - - - - Table 6 - - - - -

The number of intact pairs recalled increased as a function of the strength of pair training, as shown in Table 6 (middle column of data). Nevertheless, as the table shows in the last column of data, provided that the words had been included in the study phase of the experiment, the total number of chunks recalled (singletons plus intact pairs) remained rather constant across training conditions. The number of chunks recalled was a bit lower when the words had not been studied at all (first row of data in the table). Theoretically, it would be possible for an intact pair actually to be composed of two singletons in the mental representation that by chance happened to be recalled in the presented order, but a mathematical model indicated that this rarely happened.

In subsequent study, Chen and Cowan (in press) sought an easier way to score the results of such experiments. We also sought to determine how a limit in terms of the number of chunks recalled could be reconciled with the tendency for lists of monosyllabic words to be recalled better than lists of multisyllabic words, the word length effect of Baddeley et al. (1975). To examine these issues, pairs of monosyllabic words were taught and tested over and over, in tests in which the first member of the pair cued recall of the second. Other monosyllabic words were presented as singletons intermixed with the pairs. The entire word set was presented repeatedly until both cued recall of each pair and recognition of each singleton as such reached 100% correct. At that point, it was fair to assume that each learned pair was a single chunk and that singletons and learned pairs were equated for familiarity. (Lists of singletons that had not been included formed another control condition, and showed that familiarity training with singletons was actually not very important, probably because they were so familiar to the participants.) Both free and serial recall were used in separate experiments, and list length varied.

Opposing predictions for the proportion of words recalled correctly could be formulated on

the basis of a chunk limit or a length limit. According to a chunk limit (Cowan et al. 2004), the proportion correct should be similar for a list of n learned pairs and a list of n pre-exposed singletons because both of them contain n chunks. In contrast, according to a length limit (Baddeley et al. 1975), the proportion correct should be similar for a list of n learned pairs and a list of $2n$ pre-exposed singletons because both of them include the same number of monosyllabic words and therefore are the same length.

Table 7 shows what happened. Under some circumstances, the results conformed to the prediction based on a chunk limit almost perfectly. This happened in free recall when the lists were relatively long (6 learned pairs). The results were very similar in serial recall if the data were scored without penalty for order errors, although the table does not include that result. However, under other circumstances, the results conformed to the prediction based on a length limit almost perfectly. This happened when serial recall results were scored as correct only for items recalled in the correct serial positions, provided that the lists were relatively short (4 learned pairs). In other circumstances, intermediate results were obtained. Chen and Cowan (in press) suggested a theoretical interpretation in which there is a basic limit in the number of chunks that can be retained in working memory, but in which the core holding mechanism does not do a very good job of retaining the serial order of the chunks. To retain serial order, a phonological rehearsal mechanism (Baddeley 1986) comes into play, and it is limited to about the amount that the participant can rehearse in 2 sec (Baddeley et al. 1975). Thus, a length limit applies within the range that the phonological store can manage; that amounts to about 8 monosyllables. The serial-order-preserving mechanism might have to work in combination with the chunk-preserving mechanism.

- - - - - Table 7 - - - - -

Premise 4: The Core Working-Memory Capacity Limit is Related to the Scope of Attention

What property distinguishes between individuals who do better on tests of working memory, and also do better on aptitude tests, versus those who score lower? According to Randall Engle and his associates (e.g., Kane et al. 2001, 2004), a key characteristic is the ability to control attention. For example, Kane et al. examined performance in an *antisaccade* task in which the natural tendency to look at a suddenly-appearing object is to be resisted; the participant is rewarded for instead looking in the other direction. Similarly, others have argued that what is critical is one particular attention-related executive function, such as the ability to inhibit irrelevant information (Gernsbacher 1993; Lustig, May, & Hasher 2001; May, Hasher, & Kane 1999) or the ability to update information in working memory (Friedman, Miyake, Corley, Young, DeFries, & Hewitt in press). We (Cowan 2005; Cowan, Elliott et al. 2005) have explored the possibility that it is not one function of attention specifically, but attention more generally that is important for individual differences in working memory. The focus of attention might *zoom out* to apprehend the maximum number of items, or *zoom in* to hold on to a goal in the face of potent interference. Individuals who are better able to apprehend a field of items when the task requires it may also be the ones able to concentrate on a difficult goal when the task requires that.

In a recent study (Cowan, Fristoe, Elliott, Brunner, & Saults in preparation), we examined this contention using individual and developmental differences. A task that was assumed to reflect attention zoomed out (i.e., the scope of attention) was the two-array comparison procedure modeled after Luck and Vogel (1997). A task that was used to examine attention zoomed in to maintain a difficult goal (i.e., control of attention) was one in which there was a cue to pay attention either to a list of printed letters or to a concurrent list of spoken digits. After

each series, memory for the attended or the ignored list was tested. The measure of the control of attention was the extent to which memory for attended lists surpassed memory for ignored lists. This measure was much higher in adults than in children, who showed little evidence of being able to control their attention in this situation, even in a subsample in which children's accuracy of monitoring the attended channel was matched to the adults. Among adults, there was a significant correlation between the scope-of-attention and control-of-attention tasks, $r =$ .34. A composite measure of intelligence was related to the scope of attention, $r = .52$, and also to the control of attention, $r = .47$. The two types of attention shared 12% of the variance in intelligence; the scope of attention uniquely contributed another 15%, whereas the control of attention uniquely contributed another 10%. So there appears to be a common attention mechanism that is supplemented by some specialized components for apprehension versus control functions of attention. According to neurological evidence reviewed by Cowan (1995, 2005, in press) these would be primarily parietal versus frontal lobe mechanisms, respectively, working together closely.

In other research we have asked whether the chunk limit in working memory retention is truly a limit in attention, as we have supposed. Morey and Cowan (2004) asked whether the array-comparison procedure could be interrupted by a memory load originating in spoken digits. One would not expect such interference if visual arrays are saved entirely in an automatically-held, visually-specific buffer like the visuospatial sketch pad of Baddeley (1986). In contrast, one might expect interference if visual arrays are saved at least partly in an attention-demanding manner and attention is needed also for retention of the spoken digits. In their procedure, Morey and Cowan presented acoustic instructions for verbal recitation, followed by two arrays. After indicating whether the arrays were the same or different, the verbal stimuli (if any) were to be

recalled.   The verbal recitation conditions included (1) no recitation, (2) recitation of a random

two-digit load, which is a very light load; (3) recitation of the participant's own 7-digit telephone

number, a light load despite more digits because one's own telephone number can maintained

without holding multiple chunks in working memory; and (4) a random seven-digit memory

load.

The results are shown in Table 8. Neither a two-digit load nor recitation of the known

telephone number had much effect on visual array performance. However, a random seven-digit

load did have an effect. The effect was particularly severe when the digits were incorrectly

remembered, a situation in which attention would have been recruited to the digit-recitation task

and away from the visual arrays.

- - - - - Table 8 - - - - -

Some remaining issues were addressed by Morey and Cowan (2005). The data are shown in

some detail in Table 9, separately for different sizes of arrays. The data are shown both in term

of the proportion correct and in term of a formula that estimates the number of items retained in

working memory (Cowan 2001; Cowan, Elliott et al. 2005). The data show that the effect of a

memory load was much larger when the load was recited aloud (as was required by Morey &

Cowan 2004) than when the load was retained silently (as in other studies, such as Cocchini,

Logie, Della Sala, MacPherson, & Baddeley 2002). Aloud recitation may require more attention,

given that silent retention of verbal materials can make use of a rehearsal mechanism that

operates without much attention in adults (Guttentag 1984). The data show further that the

interference from verbal recitation occurs during the maintenance of the first array; visual array

comparison performance was impaired more when verbal recitation began between the arrays

than when it began sooner, before the first array. All of this supports the notion that visual arrays

are maintained with the assistance of attention.

- - - - - Table 9 - - - - -

Although verbal working memory benefits from rehearsal, we are working on obtaining

evidence that it, too, relies on attention in certain circumstances. One such circumstance is when

the memoranda come from several semantic categories and surpass the ability of rehearsal

mechanism, which encourages the use of semantically-based chunking mechanisms rather than

phonological rehearsal.   Bunting and Cowan (2005) presented four words in each of three

semantic categories (e.g., body parts, animals, and tools), with the words in each semantic

category presented in a different color (red, blue, or green). These were followed by a rehearsal

cue that was the semantic category (e.g., *animals?*) or a color category (e.g., *blue?*), thus

encouraging attention to the color as well as the category during presentation of the list. The

main question of the experiment was whether the presentation of the cue in a different color from

the targeted words would require extra attention. The color mismatch (second row of data in

Table 10) had a detrimental effect when the mismatch trials were rare (left-hand column of data),

but not when the mismatch trials were common (right-hand column of data). This suggests that

when mismatch trials were rare, those rare mismatch trials recruited or required attention and

interfered with recall by drawing attention away from the list items. Converging evidence on the

use of attention in verbal recall comes from a task in which running memory span is combined

with an attention-demanding, button-press task, presented during the retention interval, that is

modeled after the antisaccade task (Bunting & Cowan 2004).

- - - - - Table 10 - - - - -

Finally, we have explored other means to draw inferences about the attentional demands of

verbal working memory. We have relied upon *proactive interference*, the tendency for retrieval

to suffer interference from previous, similar material. The logic is that, if the materials to be remembered are held within the focus of attention, they are less susceptible to proactive interference than if the materials to be remembered are held at least partly in some other portion of memory, where they can be confused with previous materials (cf. Halford, Maybery, & Bain 1988). Cowan, Johnson, and Saults (2005) presented lists that were followed by a probe item. The task was to indicate as quickly as possible whether the probe item was in the list. On high-proactive-interference trials, the current trial was preceded by several other trials using items from the same semantic category, including some of the same words. Sometimes the lists were presented with all items at once and sometimes they were presented one item at a time, but the results did not differ between those conditions. When the list was short enough to fit within the focus of attention most of the time (three and four items long), there was no effect of proactive interference. However, when the list was longer (six and eight items long), there was an effect of proactive interference, as shown in Table 11. The results suggest that three and often four items were held in the focus of attention.

- - - - - Table 11 - - - - -

Other theorists have reached different conclusions. Oberauer (2002, 2005) believes that there is a 1-chunk focus of attention surrounded by a capacity-limited region or fringe holding up to about 4 chunks. However, his results also could be explained with the notion that there is a single focus of attention that holds up to about 4 chunks in the average adult, but with a prioritization of the items in the focus of attention.

Concluding Remarks

We have argued on the basis of recent evidence (most of which is summarized in Tables 1 through 11) for four related premises: (1) that working memory is not just general processing

efficiency; (2) that the relation between working memory and intellectual maturation and aptitude does not depend on using a dual task to examine working memory; (3) that working memory performance depends on the notion of capacity, expressed in terms of chunks; and (4) that the core working-memory capacity limit is related to the scope of attention. We believe it to be useful that these premises have been stated separately because evidence against one of them would not necessarily invalidate the others. Together they form a theoretical view (Cowan, Elliott et al. 2005) that has practical, theoretical, and philosophical implications.

Practically, we have shown that considering both the scope of attention and the control of attention can lead to excellent predictions of aptitude in a principled manner (Cowan et al. in preparation). These days when video games are everywhere, measurements of attention may be of higher construct validity than the nonverbal measures within intelligence tests that are used to estimate fluid or native intelligence. Those nonverbal measures were probably intended to examine how individuals deal with novel situations but the situations may be too similar to video games, television shows, or school materials to be sufficiently novel to today's test-takers.

Theoretically, we believe that we are getting closer to an understanding of the basic limits that can be used to predict performance in a range of cognitive tasks. For example, it is helpful to know when chunk limits apply and when length limits apply in verbal recall (Chen & Cowan in press). It is helpful to understand something about the attention requirements of working memory tasks (Bunting & Cowan 2005; Morey & Cowan 2004, 2005).

Philosophically, the portion of working memory that is based on the focus of attention is closely related to the conscious mind (Baars & Franklin 2003; James 1890). We hope that explorations of this aspect of working memory will lead to a more satisfying understanding of the phenomenological aspect of cognitive psychology, a bridge between behavioral results and

our personal understanding of ourselves as human beings.

<div align="center">References</div>

Baars, B.J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, **7**, 166-172.

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

Baddeley, A., & Hitch, G.J. (1974). Working memory. In G. Bower (ed.), *Recent advances in learning and motivation*, Vol. VIII. New York: Academic Press.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (eds.), *Models of Working Memory: Mechanisms of active maintenance and executive control*. Cambridge, U.K.: Cambridge University Press. (pp. 28-61)

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**, 575-589.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, **133**, 83-100.

Broadbent, D.E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (eds.), *Studies in long-term memory*. Wiley. (pp. 3-18)

Bunting, M.F., & Cowan, N. (2004, November). *Working-memory retrieval takes attention: Effects of distraction under time pressure*. Poster presented at the annual meeting of the Psychonomic Society, Minneapolis.

Bunting, M.F., & Cowan, N. (2005). Working memory and flexibility in awareness and

attention. *Psychological Research*, **69**, 412-419.

Case, R., Kurland, D.M., & Goldberg, J. (1982). Operational efficiency and the growth of short-
     term memory span. *Journal of Experimental Child Psychology*, **33**, 386_404.

Chen, Z., & Cowan, N. (in press). Chunk limits and length limits in immediate recall: A
     reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Cocchini, G., Logie, R.H., Della Sala, S., MacPherson, S.E., & Baddeley, A.D. (2002).
     Concurrent performance of two memory tasks: Evidence for domain_specific working
     memory systems. *Memory & Cognition*, **30**, 1086_1095.

Conlin, J. A., Gathercole, S. E., & Adams, J. W. (2005). Children's working memory:
     Investigating performance limitations in complex span tasks. *Journal of Experimental Child
     Psychology*, **90**, 303_317.

Conway, A.R.A., Cowan, N., Bunting, M.F., Therriault, D.J., & Minkoff, S.R.B. (2002). A
     latent variable analysis of working memory capacity, short-term memory capacity,
     processing speed, and general fluid intelligence. *Intelligence*, **30**, 163-183.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W.
     (in press). Working memory span tasks: A methodological review & user's guide.
     Psychonomic Bulletin & Review.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology
     Series, No. 26. New York: Oxford University Press. (Paperback edition: 1997)

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental
     storage capacity. *Behavioral and Brain Sciences*, **24**, 87-185.

Cowan, N. (2005). Working-memory capacity limits in a theoretical context. In C. Izawa & N.
     Ohta (eds.), *Human learning and memory: Advances In theory and applications*. The 4th

Tsukuba international conference on memory. Erlbaum. (pp. 155-175)

Cowan, N. (in press). *Working memory capacity*. New York, NY: Psychology Press.

Cowan, N., Chen, Z., & Rouder, J.N. (2004). Constant capacity in an immediate serial-recall

task: A logical sequel to Miller (1956). *Psychological Science*, **15**, 634-640.

Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A., & Conway,

A.R.A. (2005). On the capacity of attention: Its estimation and its role in working memory

and cognitive aptitudes. *Cognitive Psychology*, **51**, 42-100.

Cowan, N., Elliott, E.M., Saults, J.S., Nugent, L.D., Bomb, P., & Hismjatullina, A. (in press).

Rethinking speed theories of cognitive development: Increasing the rate of recall without

affecting accuracy. *Psychological Science*.

Cowan, N., Fristoe, N.M., Elliott, E.M., Brunner, R.P., & Saults, J.S. (in preparation). *Scope of

attention, control of attention, and intelligence in children and adults*.

Cowan, N., Johnson, T.D., & Saults, J.S. (2005). Capacity limits in list item recognition:

Evidence from proactive interference. *Memory*, **13**, 293-299.

Cowan, N., Nugent, L.D., Elliott, E.M., Ponomarev, I., & Saults, J.S. (1999). The role of

attention in the development of short-term memory: Age differences in the verbal span of

apprehension. *Child Development*, **70**, 1082-1097.

Cowan, N., Wood, N.L., Wood, P.K., Keller, T.A., Nugent, L.D., & Keller, C.V. (1998) . Two

separate verbal processing rates contributing to short-term memory span. *Journal of

Experimental Psychology: General*, **127**, 141-160.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and

reading. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450-466.

Daneman, M., & Merikle, P.M. (1996). Working memory and language comprehension: A

Meta-Analysis. *Psychonomic Bulletin & Review*, **3**, 422-433.

Duff, S.C., & Logie, R.H. (2001). Processing and storage in working memory span. *Quarterly Journal of Experimental Psychology*, **54(A),** 31-48.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: Genera*l, **128**, 309-331.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language*, **51**, 136–158.

Friedman, N.P., Miyake, A., Corley, R.P., Young, S.E., DeFries, J.C., & Hewitt, J.K. (In press). Not All Executive Functions Are Related to Intelligence. *Psychological Science*.

Gathercole, S.E., & Hitch, G.J. (1993). Developmental changes in short_term memory: A revised working memory perspective. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds.), *Theories of memory*. Hove, England, UK: Lawrence Erlbaum Associates. (pp. 189_209)

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

Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. *Journal of Verbal Learning & Verbal Behavior*, **13**, 114-131.

Gobet, F., & Simon, H.A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, **6**, 225-255.

Guttentag, R. E. (1984). The mental effort requirement of cumulative rehearsal: A developmental study. *Journal of Experimental Child Psychology*, **37**, 92_106.

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.

Hitch, G.J., Towse, J.N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, **130**, 184-198.

Hockey, R. (1973). Rate of presentation in running memory and direct manipulation of input-processing strategies. *Quarterly Journal of Experimental Psychology*, **25(A),** 104-111.

James, W. (1890). *The principles of psychology*. NY: Henry Holt.

Johnson, N.F. (1978). The memorial structure of organized sequences. *Memory & Cognition*, **6**, 233-239.

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

Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, **130**, 169-183.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. E. (2004). The generality of working_memory capacity: A latent_variable approach to verbal and visuo_spatial memory span and reasoning. *Journal of Experimental Psychology: General*, **133**, 189-217.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, **14**, 389-433.

Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high level cognition ? *Psychonomic Bulletin & Review*, **12**, 165-170.

Logie, R.H., Cocchini, G., Della Sala, S., & Baddeley, A.D. (2004). Is there a specific executive

capacity for dual task co_ordination ? Evidence from Alzheimer's Disease. *Neuropsychology*, **18**, 504-513..

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, **390**, 279-281.

Lustig, C., May, C.P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, **130**, 199-207.

May, C.P., Hasher, L., & Kane, M.J. (1999). The role of interference in memory span. *Memory & Cognition*, **27**, 759-767.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.

Marmurek, H.H., & Johnson, N.F. (1978). Hierarchical organization as a determinant of sequential learning. *Memory & Cognition*, **6**, 240-245.

Morey, C.C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, **11**, 296-301.

Morey, C.C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 703-713.

Murray, D.J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, **78**, 679-684.

Pollack, I., Johnson, I.B., & Knaff, P.R. (1959). Running memory span. *Journal of Experimental Psychology*, **57**, 137-146.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **28**,

411-421.

Oberauer, K. (2005). Control of the contents of working memory—A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 714–728.

Pylyshyn, Z.W., & Storm, R.W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179-197.

Ryan, J. (1969). Grouping and short-term memory: Different means and patterns of groups. *Quarterly Journal of Experimental Psychology*, **21**, 137-147.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74** (Whole No. 498.)

Tulving, E., & Patkau, J.E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology*, **16**, 83-95.

Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, **28**, 127-154.

Table 1

<u>Findings of Cowan et al. (1998), Experiment 1</u>

Increase in span resulting from a 1-SD speedup in retrieval rate (path coefficient): .41 SD

Increase in span resulting from a 1-SD speedup in rehearsal rate (path coefficient): .49 SD

The proportion of the total variance accounted for by these two rates: 60%

The portion of within-age variance in span accounted for by these two rates: 87%

Table 2

Findings (Retrieval rate; Span) of Cowan et al. (in press), Experiment 2

Group and Experimental Phase	Retrieval rate	Span
Training-group children, non-speeded Phase 1:	1.10 items / sec;	5.37 items
Training-group children, speeded Phase 2:	2.19 items / sec;	5.37 items
Control-group children, non-speeded Phase 1:	1.16 items / sec;	5.53 items
Control-group children, non-speeded Phase 2:	1.41 items / sec;	5.63 items
Adult comparison group, non-speeded	1.63 items / sec;	7.56 items

Table 3

Response Durations In Seconds For Two-Item Lists, After Cowan et al.(2003), Experiment 2

_____

Measure	Grade 3	Grade 5	Adult
Digit span:	1.57	1.64	1.40
Counting span:	2.43	1.94	1.58
Listening span:	5.74	3.74	2.48

_____

Table 4

Proportion of the Within-Age-Group Variance in *g* That Is Accounted for by Three Types of

Tasks in Two Age Ranges (after Cowan, Elliott et al., 2005)

_____

Type of Working-Memory Task	Grades 2, 4	Grade 6, College
Traditional storage-and-processing dual tasks	.15*	.15*
No dual task, but rehearsal is not feasible	.14*	.11*
Digit span; rehearsal is possible for older participants	.14*	.01, n.s.
All working-memory measures taken together	.24*	.17*

_____

* p < .05, regression

Table 5

Some Procedures in which Each Item Remains a Separate Chunk (after Cowan, 2001)

_____

Measuring the number of stimuli that are recalled perfectly, which would not be expected if

  recall was based on a chunking or rehearsal strategy (Broadbent, 1975)

Noting the way in which long-term retrieval depends on repeatedly refilling working memory

  capacity, so that bursts of items are produced (Broadbent, 1975)

Imposing information overload in a spatial array to be remembered (Sperling, 1960)

Imposing information overload in a rapid and unpredictable series to be remembered, as in

  running memory span (Hockey, 1973)

Requiring memory of a verbal sequence that was ignored at the time of its presentation (Cowan,

  Nugent, Elliott, Ponomarev, & Saults, 1999)

Imposing a rehearsal-prevention task along with a verbal series (Murray, 1968)

Using verbal material in which the chunks are too long to rehearse (Glanzer & Razel, 1974)

Using materials in which the configuration of items keeps changing, as in multi-object tracking

  (Pylyshyn & Storm, 1988).

Examining elements that were used within a newly-formed chunk (Ryan, 1969)

_____

Table 6

Recall Statistics from Cowan et al. (2004), Experiment 1

Type of Study	Chunks Recalled per List		
	Singletons	Intact Pairs	Total
Words studied as singletons 0 times, pairs 0 times	1.44	1.40	2.83
Words studied as singletons 4 times, pairs 0 times	1.79	1.54	3.33
Words studied as singletons 1 times, pairs 3 times	1.25	2.13	3.38
Words studied as singletons 2 times, pairs 2 times	1.27	2.21	3.48
Words studied as singletons 0 times, pairs 4 times	0.46	3.04	3.50

Table 7

Proportion of Words Correctly Recalled in Lists by Condition, Chen and Cowan (in press)

_____

Type of Recall	Prior Exposure to Words As	Proportion of Words Correct	Learned Pairs: Limiting Factor(s)
Free Recall	4 Learned Pairs	0.84	Both factors?
Free Recall	4 Singletons	0.93	
Free Recall	8 Singletons	0.60	
Free Recall	6 Learned Pairs	0.73 [a]	Chunk limit
Free Recall	6 Singletons	0.75 [a]	
Free Recall	12 Singletons	0.45	
Serial Recall	4 Learned Pairs	0.54 [b]	Length limit
Serial Recall	4 Singletons	0.95	
Serial Recall	8 Singletons	0.48 [b]	
Serial Recall	6 Learned Pairs	0.35	Both factors?
Serial Recall	6 Singletons	0.54	
Serial Recall	12 Singletons	0.18	

_____

Note. Proportions with the same letter superscript do not significantly differ.

Table 8

Proportion of Correct Comparison of Two Visual Arrays Under Various Digit Recital Conditions (from Morey & Cowan, 2004), Collapsed Across 4-, 6-, and 8-Item Arrays

Memory Load Condition	Proportion Correct
No memory load	.91
Two-digit memory load	.92
Own seven-digit phone number recited	.90
Seven-digit memory load recited (all trials)	.85
Seven-digit memory load incorrectly recalled	.74

Table 9

Proportion Correct Comparison of Two Visual Arrays (and Capacity Estimate) Under Various

Array and Digit Load Conditions, After Morey & Cowan (2005), Experiment 1

_____

	Number of Items Per Array		
Verbal Task Condition	4	6	8
No verbal task	.95 (3.60)	.87 (4.40)	.80 (4.80)
Seven-digit load silently held	.93 (3.44)	.85 (4.20)	.75 (4.00)
Seven-digit load recited from before Array 1	.85 (2.80)	.75 (3.00)	.65 (2.40)
Seven-digit load recited between arrays only	.79 (2.32)	.69 (2.28)	.62 (1.92)

_____

Table 10

Proportion of Targeted Words Recalled in Various Retrieval-Cue Conditions, After Bunting and

Cowan (2005)

_____

	Percentage Mismatch Trials	
Type of Retrieval Cue	12.5%	50%
Category cue presented in same color as targeted items	.62	.55
Category cue presented in a color different from targets	.54	.56
Category cue presented in a neutral, black color	.63	.56

_____

Note. Color-cue trials (not shown) accounted for half of the different-color (i.e., mismatch) trials

in the 50%-mismatch condition, but none in the 12.5%-mismatch condition.

Table 11

Response Times for Different Proactive Interference Conditions and Set Sizes, After Cowan,
Johnson, and Saults (2005)

	Proactive Interference Condition (milliseconds)		
List Length	Low	High	Difference
Three Items	851	853	2, n.s.
Four Items	906	923	17, n.s.
Six Items	940	994	54 *
Eight Items	965	1058	93 *

---

* $p < .05$, Tukey Test