THEORY AND MEASUREMENT OF WORKING MEMORY CAPACITY LIMITS

Nelson Cowan, Candice C. Morey, Zhijian Chen,
Amanda L. Gilchrist, and J. Scott Saults

Abstract

We review the evidence for various kinds of limit in the capability of working memory, the small amount of information that can be held in mind at once. To distinguish between types of limit in working memory, we invoke metaphors of space (capacity), time (decay and speed), and energy (control of attention). The review focuses primarily on recent evidence on a limit in how many chunks can be held in working memory, how this kind of limit can be measured, and how it can be distinguished from other types of limits. We explore the theoretical and practical importance of different working memory limits in research that is nomothetic (referring to general laws) and ideographic (referring to individual and group differences). The appropriate measure of working memory depends on one's holistic or analytic scientific interest. © 2008, Elsevier Inc.

THE PSYCHOLOGY OF LEARNING
AND MOTIVATION VOL. 49
DOI: 10.1016/S0079-7421(08)00002-9

0079-7421/08 $35.00

I. Key Theoretical Issues

A. Introduction

Cognitive psychology casts a spotlight on both what people can do and what they cannot do. A person often can keep thinking about an important goal for some time, as when a driver knows he or she must watch for the second right turn. However, the driver can do so only by forfeiting other processing, such as complex conversation with a passenger. Watching a soccer game, a viewer can observe several players on the field at once but often will experience surprising (to the viewer) lapses in awareness of what is going on elsewhere on the field or in the stadium at that moment.

In describing the miracle of what humans can do, one inevitably is describing also the limits to what they can do. Although various animals including humans can visually scan an entire field or forest at once looking for predators or prey, humans excel at sequestering a small portion of the information to allow amazingly in-depth analysis of the selected portion. This is what the research literature shows. Miller (1956) wrote his famous article about humans typically being limited to remembering about seven items at once and, shortly afterward, Miller, Galanter, & Pribram (1960) wrote about how this limited memory may act as a "working memory" to keep in mind goals and other information that one needs to complete a task. In the present book series, Baddeley and Hitch (1974) wrote a seminal article on working memory suggesting that it must be composed of several parts that operate separately. Since then, research on the topic of working memory certainly has blossomed, both in the form of behavioral research on the topic and, especially recently, in the form of related neurobiological research. In our chapter, we focus on working memory capacity limits, with special attention on item limits like Miller (1956). To keep things simple we will not go into the various other terms that are used for similar concepts of a temporary memory, including short-term memory and immediate memory. These terms will be used interchangeably with working memory.

The topic of working memory limits is quite broad, considering that working memory is involved in almost every cognitive task and often sets boundaries for the performance of that task. For example, one cannot successfully complete arithmetic problems without some working memory of intermediate results and what types of calculation are still to be done. Probably for this kind of reason, researchers have approached the topic of working memory limits from a number of different theoretical vantage points. It is important to acknowledge these vantage points in order to avoid being stuck thinking that different investigators disagree on substantive points when, in many cases, they simply are interested in different issues.

Three issues that we introduce below refer to the type of working memory limits, the goal in studying these limits, and the level of analysis at which the limits are studied. We explain where our own recent research fits into this family of questions and what it is telling us.

The main emphasis of the present review is on the possibility of quantifying and characterizing limits in the number of chunks of information that individuals can retain in working memory at once, the theoretical and practical significance of this chunk capacity limit, and how it can be distinguished from other types of limits on working memory. We do not go into great detail regarding the specifics of different theoretical frameworks but we do have much in common with theorists who, in various ways, have emphasized a key role of attention in understanding the strengths and limits of both working memory storage and information processing (e.g., Awh & Jonides, 2001; Baddeley, 2000; Cowan, 1995, 1999, 2001; Davelaar, Goshen-Gottstein, Ashkenazi, Haarman, & Usher, 2005; Grossberg, 1978; Lovett, Reder, & Lebière, 1999; Unsworth & Engle, 2007).

B. Different Types of Working Memory Limits

After Salthouse (1985; also Kail & Salthouse, 1994), we point to an analogy between limits in working memory and limits in physical events, which occur within a certain time and space and involve a certain amount of energy. Working memory representations might be *limited in time*; they could fade quickly over time even in the absence of any sort of interference. Alternatively, working memory representations might last over time, but only until they are displaced by other representations that become active because of outside events or internal thoughts. The notion (Miller, 1956) that only a certain number of items can be held at once is like a space limit in which, say, only a certain number of eggs can fit in an egg carton. This is what we will term *chunk capacity limits*. The third possibility is that there is an energy limit, in which electrophysiological activity is in some not-quite-defined sense a type of energy. If the representation of each item required a certain amount of this neural energy per unit of time and other mental processes did as well, then any given representation would face competition from other representations in working memory or other mental processing that used the energy. This type of limit is often referred to as *resource limits*. This tripartite taxonomy is not meant as an assertion that no other factors can influence the fate of a representation in working memory. For one thing, multiple limits may apply (e.g., both space and energy). Moreover, there are additional factors. For example, the ability to use knowledge to form larger chunks of information eases the load on working memory (Miller, 1956). Thus, the letter sequence *irsciafbi* is much easier to remember in working memory if one notices that it is composed of three-letter acronyms for US government agencies, the Internal Revenue Service, the Central Intelligence Agency, and the Federal Bureau of Investigation. The present chapter is focused on identifying space or chunk capacity limits, and attempting to specify the relation between these types of limits and possible limits in time and energy.

C. Nomothetic and Ideographic Questions

Nomothetic questions relate to how a process normally or typically operates, whereas ideographic questions relate to how individuals (or groups) differ from this norm. Both of them are important and have played an important role in research on working memory. Perhaps most importantly, they should not be confused with one another.

Much of the early research on working memory was nomothetic in nature. Miller (1956) noted that the normal working memory limit was about seven items (give or take two) but he did not have much information on ideographic patterns. A large literature grew up following Baddeley and Hitch (1974) in which various manipulations were used to isolate components of working memory in the normal individual (Baddeley, 1986; Cowan, 2005; Gathercole & Baddeley, 1993; Logie & Gilhooly, 1998; Shah & Miyake, 2005). For instance, a phonological distracting task would interfere with verbal memory, but not spatial memory, whereas a spatial interfering task would do the reverse. A simple distracting task, such as repeating one letter would interfere with automatically held verbal memory, whereas only a more complex and engaging task, such as generating random numbers, would interfere with the central executive processes thought to be needed to control the flow of information from one memory store to another (Baddeley, 1996).

There also were ideographic approaches early on. The invention of the digit span task, in which the experimenter estimates how many digits the subject can repeat, was developed to help determine the maturational age equivalent of a particular child (e.g., Binet & Simon, 1916/1980; Bolton, 1892; Jacobs, 1887; Wechsler, 1944, 1991). That is, clearly the normal digit span increased with age in childhood and the question was the age norm that matched performance in a particular child. However, information carried over from nomothetic experimental methods can be of great help in solving ideographic questions. For example, the testing techniques developed by Alan Baddeley and colleagues have been used in many studies by Susan Gathercole and colleagues to study how different components of working memory change with age and how individual differences in each component are related to learning disabilities (for a review, see Baddeley, Gathercole, & Papagno, 1998).

It has also been tried in the other direction: ideographic differences have been used to try to help identify nomothetic principles. A good example is a study by Baddeley and Warrington (1970) on memory in amnesia. Informal study of amnesic individuals showed that they had great difficulty learning new information but still could interact rather normally (according to casual observation, at least) so that it seemed likely that they had preserved short-term memory. Baddeley and Warrington used an experimental method on amnesic individuals to verify that the typically nomothetic use of this same method provided separate measures of short- and long-term memory. In particular, it had been theorized that, in free recall, the most recent list items include a strong contribution of short-term memory, whereas earlier portions of the list are recalled based on long-term memory. Strongly supporting this division between short- and long-term memory and the measures of them in free recall, amnesic individuals were deficient in recall of earlier portions of the list but not in the last few list items. In more recent work, Engle and colleagues have shown that individuals with relatively low working memory span scores tend to be deficient in the control of attention compared to high-span individuals. They also have shown that high-span individuals with distraction perform similarly to low-span individuals without distraction (Kane & Engle, 2000; Rosen & Engle, 1997), and they have used that information to suggest that a key component of the normal operation of working memory is, in fact, the control of attention.

The use of ideographic differences to help identify nomothetic principles works fine to a point, but there are limits to its application. First, it is sometimes found that not all sources of individual differences converge on the same mechanisms. For example, Cowan et al. (1998) found that the effects of age on the timing of digit span recall were completely different from the effects of individual differences within an age group. If different sources of ideographic information yield different interpretations then one cannot tell which of them reveals the most important aspect of normal functioning. For example, according to one study on rapidly perceiving several items at once, or *subitizing*, there is no important individual difference in the number of items that can be held in working memory at once, in contrast to the pronounced individual differences in the control of attention (Tuholski, Engle, & Baylis, 2001). From this, one might reason that thinking of working memory as a capability of holding a certain number of items is the "wrong way" to think about working memory. However, other ideographic sources challenge this conclusion. In Section III.D, we describe a study by Basak and Verhaeghen (2003) that shows the subitizing range to change with adult aging, in contrast to the finding of Tuholski et al. (2001) in the examination of individual differences in young adults. Some studies have, indeed, used age differences in the holding capacity of working memory to explain developmental differences in cognition (Andrews & Halford, 2002; Pascual-Leone, 1970, 2005). In Section III.E, we will illustrate the importance of capacity with a procedure by Gold et al. (2006) that shows how space and energy constraints on working memory can be separated, strongly supporting space (i.e., chunk capacity) differences between normal and schizophrenic individuals; and we will examine a number of related examples.

A second limit to the research strategy of using ideographic information to discern nomothetic principles is that the latter can be theoretically important even without producing individual differences. To draw an analogy, most individuals do not differ in how many limbs they have (4), but the number of limbs is an important aspect of human physiology. The number of items retained in working memory at once is less visible than the number of limbs, but it similarly would be important even if it did not differ among individuals. Knowing that humans can retain 3 or 4 separate pieces of information at once, or can form 3 or 4 relations at once, allows many predictions about task performance that would not be possible otherwise (e.g., Halford, Baker, McCredden, & Bain, 2005; Halford, Cowan, & Andrews, 2007).

A key, capacity-limited part of working memory may comprise the contents of the focus of attention (Cowan, 2001). If so, it is not yet clear whether attention contents and attention control depend on separate mechanisms or a common mechanism. Here again, though, the ideographic information can be over-interpreted. Suppose it were found that there is a very high correlation between the control of attention and the holding capacity of working memory. Then it might be concluded that both of them stem from a single mechanism. For example, it might be that holding capacity depends on the ability to filter out extraneous, irrelevant items (Hasher & Zacks, 1988; Vogel, McCollough, & Machizawa, 2005). However, it would be going beyond the evidence to conclude from that that the control of attention and the holding capacity are not separate mechanisms, each with its own nomothetic importance. By analogy, the correlation of individuals' arm and leg length is rather high, but still one would not conclude that the arms and legs are indistinguishable. Arms and legs are independently important for overall active motor function, even if they strongly covary. As another example, having two eyes is important even though individuals do not vary in the number of eyes. Thus, one cannot use the ideographic uniformity of a component to rule out the nomothetic importance of that component. Thus, a particular measure of holding capacity would be nomothetically important even if it did not vary among individuals or age groups.

In our laboratory we have taken an approach that embraces both nomothetic and ideographic evidence, allowing cross-fertilization between the two while staying mindful that both approaches are meaningful in their own right. We will present our thoughts and our evidence on working memory limits that could be classified as space limits (i.e., chunk capacity limits), energy limits (i.e., the limits of total complexity and of attention control), and time limits (i.e., temporal decay of memory). These thoughts are tentative as the understanding of working memory limits is still evolving. Yet, we believe that considerable progress is being made by ourselves and others.

D. Holistic and Analytic Approaches

There are different attitudes toward research in psychology that one can term holistic and analytic. The holistic approach may be clearest to those interested in practical applications but they won't get far without including an analytic approach as well, in our opinion. Let us suppose that working memory is composed of multiple processes and components operating together, as Baddeley and Hitch (1974) assumed. (For now, it does not matter whether the components are just as Baddeley and Hitch thought.) Then a holistic approach is one in which we ask what is the combined effect of all components on working memory. In contrast, an analytic approach is one in which we ask what is the effect of each component. The components would have to be experimentally isolated to answer that question. If one adopts a holistic approach, it is not clear if good evidence could be found for time, space, or energy limits in working memory. That is because the components might operate differently and compensate for one another's weaknesses. If one sort of memory representation is lost as a function of time (i.e., decays), another working memory function might serve to counteract decay so that one cannot see it. A classic example is the use of covert verbal rehearsal to refresh representations that presumably would decay if they were not refreshed (Baddeley, 1986). If one sort of memory representation has a space limitation, another sort may have no such limitation so that the combined effect is to obscure the limit. If there is an energy limit, some strategies might have the effect of overcoming that limit too. So, as we will see, steps must be taken to examine each component in isolation. For example, assuming that one cannot rehearse memoranda while repeating a different word, it is possible to see whether memory is lost as a function of time during that repetition (Baddeley, 1986). (It turns out that the question of whether one can see decay in such a circumstance is still not satisfactorily resolved.) The issue of holistic and analytic approaches is a complex one to be revisited in Section IV.

II. Space (Chunk Capacity) Limits

A. A Brief History of the Capacity Concept

We refer the reader to Cowan (2005) for a more thorough exposition of the history of the capacity concept. Miller (1956) set the world on fire with his thorough demonstration that there are some real limits on how long a series of random items, such as words or digits, can be before people are unable to repeat the series (about 7 items). In the same paper, he demonstrated some similar limits on how many simple items, such as tones of different frequencies, can be included in an absolute identification test before people are unable to identify individual items (about 7 items), and how many items can be included in a set to be rapidly enumerated before it becomes necessary to count them one by one (about 7 items, he thought; actually it is closer to 4 items as in the subsequent literature reviewed by Cowan, 2005). Without making a strong theoretical claim, Miller provided evidence that inspired many researchers to imagine that there might be a very general mental resource with an easily quantifiable limit, and that the limit might be used to predict what materials are easily perceived, understood, or remembered, and what materials are so processed only with difficulty. Shortly afterward, when Broadbent (1958) published his famous book on information processing, it included a simple processing model in which information was dealt with in an effective manner only if it went through a limited-capacity system. It was natural to suppose that the limits such as those that Miller pointed out applied to the limited-capacity system that Broadbent proposed. A good metaphor for the limited-capacity processor therefore seemed to be a box that can hold only a certain, limited number of items of particular type because of its limited internal space (like an egg carton). That metaphor was advanced further by Atkinson and Shiffrin (1968), who reported in their book chapter some sophisticated empirical and modeling work related to storage capacity limits. Miller et al. (1960) discussed how this limited-capacity system for storing information temporarily was essential to keep in mind goals as one plans and carries out activities, and thus referred to it as working memory.

The storage limit per se was not the main contribution of Miller (1956). A more critical contribution was his noticing what type of measure one had to use to observe near-constant capacity. In the zeitgeist of that time, when general-purpose computers had recently been invented, it had been assumed that communicating devices impart information that is to be measured in terms of the number of choices that could have been made, called binary digits, or bits. Within a string of digits in which each one ranges from 0 to 9, there are 10 choices, which require between 3 and 4 binary decisions. (For example, one could often, though not always, guess a digit after only three yes-or-no questions. One could ask, is the digit less than 6? If so, is it less than 3? If not, is it less than 5? If so, it must be either 3 or 4. N binary decisions, reflecting N bits, allow perfect selection from 2^N choices, and there are ways to calculate the bit size of any number of choices, many of which require fractions.) By this metric, it should be much easier to recall digits, which come from a set of 10, than words, which come from a set of many thousands. Yet, Miller saw that people remember almost as many words as digits. For reasons like this, he proposed that the limit in capacity was to be measured not in bits, but in whatever units are psychologically meaningful, which he called chunks. He also explained how the process of chunk formation, or chunking, could be used to reduce the load on memory. A good example is the invention of the name ROY G. BIV to encode easily the seven colors of the rainbow: red, orange, yellow, green, blue, indigo, and violet.

From this point on, though, the situation became more complicated. A chapter by Baddeley and Hitch (1974) in this book series provided reasons why a single, general capacity-limited faculty would not suffice to account for the various evidences. Interference with verbal materials comes from other verbal materials and interference with nonverbal, spatial, or visual materials comes from other such materials, leading to the suggestion that there are separate verbal and visual/spatial types of active memory representations. Baddeley and Hitch allowed that there could be a central type of memory representation also, but Baddeley removed it from later versions of his model (e.g., Baddeley, 1986) and restored it only recently with the inclusion of the episodic buffer (Baddeley, 2000, 2001). Baddeley and Hitch also reframed the issue of temporary memory by emphasizing that it was essential in various types of higher-level cognition such as reasoning, and thus referred to the ensemble as working memory. In the model that evolved (e.g., Baddeley, 1986), working memory included low-level phonological and visual–spatial stores with a short time limit, managed by central executive processes.

Broadbent (1975) also recognized the contribution to working memory of processes other than a central, capacity-limited store, which, for example, could include the formation of multi-item chunks on line, other types of rapid memorization, and contributions of rehearsal. He therefore suggested that Miller's (1956) "magical number seven" was the result of an ensemble of processes, one of which is a central, capacity-limited store of not seven items, but three. One could observe it in certain restricted circumstances: when one looks at the number of items that can be repeated not half the time, but almost always without error, and when one looks at how many items are recalled from a category in long-term memory in a single, rapid burst. For example, a person trying to recall aloud as many countries as possible tends to recall them in small bursts such as Venezuela–Brazil–Columbia, United States–Canada–Mexico, and England–France.

Cowan (2001) elaborated upon Broadbent's (1975) approach by documenting many more situations in which it seemed unlikely that multi-item chunks could be formed and seemed likely that each presented item was a single chunk. In such situations, participants appear able to retain and recollect 3 to 5 items in immediate memory tests. Examples include many studies in which a random list of familiar words was presented acoustically, so that each word would be fully encoded into a single unit, and the participant was to carry out articulatory suppression (repetition of a single word over and over during the presentation of the list, in this case not loudly enough to prevent this encoding), so that verbal rehearsal could not be used to refresh the word representations or link words together to form higher-level chunks. In other examples, verbal lists were presented in an unattended channel within a selective attention task, or simple visual objects were presented all at once, both of which discourage rehearsal and chunking. A wide variety of procedures seemed to converge on an estimate of 3–5 familiar units that could be retained at the same time without further chunking.

What the approach of Broadbent (1975) and Cowan (2001) illustrates is an analytic view of working memory capacity limits, in contrast with the more holistic view of Miller (1956). They are not making the assertion that one can find this limit in every immediate memory task. Rather, the assertion is that if one analyzes tasks, one finds multiple mechanisms and one can isolate and thereby identify a mechanism that can hold just a few chunks at a time. Cowan (2001, 2005) suggested that the chunk-capacity-limited mechanism of working memory is the focus of attention. Oberauer (2002, 2005) took a similar view but suggested that there is a focus of attention that holds only 1 item and a fringe that holds the 3–5 items suggested by Cowan (2001). Cowan (2005) alternatively suggested that his evidence could be explained if there is a larger focus of attention that holds 3–5 chunks, but with some chunks at a higher priority than others. The resources of the focus of attention presumably must be divided among chunks, but cannot be divided among more than a handful of these chunks (with a range possibly spreading between 2 and 6 chunks in normal adults).

One special use of the focus of attention would be to form new chunks, which would require that the items to be associated or bound together be in the focus of attention at the same time. Thus, there would be a processing limit but any amount can, in principle, be remembered if enough time is taken to form higher-level chunks. An expert at forming chunks from a certain type of material can increase working memory substantially; several individuals have learned to increase their digit spans from the normal value of about 7 items to 80 or more (Ericsson, Chase, & Faloon, 1980). This acquired skill did not generalize to other materials such as letters. We assume that the fundamental capacities remain the same but that the size of chunks that mnemonists recall has increased enormously for the types of materials used in training (see also Ericsson, Delaney, Weaver, & Mahadevan, 2004).

In sum, holistically people typically can remember about 7 items but the value varies a great deal according to circumstances and training. Analytically, we believe that we can find evidence of a capacity limit of about 3 to 5 chunks, but probably with considerable individual variation (which we will discuss later). Other memory processes typically cloud the observation of a chunk capacity limit.

Instead of repeating in more detail the evidence from Cowan (2001) on the capacity limit as observed when the formation of multi-item chunks is prevented, below we search for various means of examining and controlling the chunking process and other processes that may cloud our observation of capacity limits. The basic hypothesis under investigation is that the same fundamental chunk capacity limit will hold in these more complex circumstances.

B. Further Evidence of Verbal Chunk Capacity Limits

Various attempts were made following Miller (1956) to determine how many chunks have been formed (e.g., Glanzer & Razel, 1974; Marmurek & Johnson, 1978; Slak, 1970; Tulving & Patkau, 1962). In what appears to be the first and perhaps the most relevant of these, Tulving and Patkau (1962) presented word lists following sequences with seven levels of approximation to English, for immediate free recall (i.e., with no constraint on the order in which the words should be recalled). The levels of approximation were formed by varying the number of words in a row that were syntactically and semantically reasonable according to participants who read them. For example, in one level, participants received two words that already had been judged to go together (e.g., *man thought*) and had to add a third word (e.g., *that*) whereas another participant was then to continue the process (e.g., *thought that* ___, which might be filled in with *way*, leading to the total sequence *man thought that way* ___, and so on). These associations could be used to lessen the load on working memory through the meaningful constraints of language. A medium-low level of approximation could be a sequence like, *Man thought that way we go home now or never think badly of him and us as*. Notice that although short subseries of words go together, the entire sentence is nonsensical and syntactically incorrect. The results from immediate recall of series with various levels of approximation to English were all scored according to a simple assumption of what forms a chunk. The assumption was that words within a chunk would be recalled in immediate sequence, whereas there would be no sequential succession for words recalled from different chunks. In the hypothetical example above, if the participant's recall protocol was *we go home never think badly thought that as* it would thus be scored as including four of what were called adopted chunks: *we go home*, *never think badly*, *thought that*, and *as*. It was found that, in every condition, participants recalled about 4–6 adopted chunks of information. The increasing levels of approximation to English simply resulted in larger chunks being recalled, not more chunks. This is what would be expected if there is a basic chunk limit but the size of chunks depends on the participant's knowledge of the materials.

The interpretation of Tulving and Patkau (1962) is similar to that of Ericsson et al. (1980, 2004) based on participants who learn to repeat long sequences of digits by forming large chunks, starting with multi-digit values they already knew as sports enthusiasts (e.g., 3:59, Roger Bannister's record when he broke the 4-min mile). In the latter case, it also was possible to learn to organize several chunks into higher-order chunks. Still, a basic capacity of just several chunks could be assumed.

Baddeley, Thomson, and Buchanan (1975) set off a firestorm that seemed to go against this finding that there is a fixed chunk capacity limit. They showed that, under some circumstances at least, a list of short (e.g., monosyllabic) words was recalled better than a list of the same number of longer (e.g., polysyllabic) words. They also found that an individual's recall was equal to the number of items that he or she could pronounce in about 2 s. This led to a theory that the limit in memory was not the number of chunks, but the amount of time for which the materials had to be remembered. The theory was that items were refreshed by rehearsal in a repeating loop (the *phonological loop* in the model of Baddeley, 1986). If a given item was not refreshed in about 2 s, it was lost from the working memory representation through temporal decay. These data were obtained using a procedure in which the task was serial recall (recall of the list items in order), the results were scored correct only for items recalled in the correct serial positions, and items were drawn from a small set used over and over throughout the experiment.

We will return to this word length effect of Baddeley and colleagues in the section on time limits. For now, what is notable is that the rehearsal strategy and scoring methods that yield this pattern of results is rather specific. If the materials are so long that using a repeating rehearsal method is impractical, or if free recall is used, participants seem less likely to apply a rehearsal strategy. Glanzer and Razel (1974) presented lists of 15 known proverbs for free recall. In that circumstance, one can see from a figure in their article, by adding the proportions across serial positions, that about 5.5 proverbs were recalled. Glanzer and Razel offered a smaller number of 2.2 by including only items recalled well because of their placement at the end of the list, assuming only those items to be part of a short-term store. Another way to look at the situation, however, is that most items have to be encoded in a form that allows them to be resistant to interference from subsequent items or from the process of recall. The last few items can be recalled immediately, in which case they are not susceptible to that interference and do not have to be encoded into working memory as thoroughly. (Recall might occur, e.g., from sensory memory.) If one replaces the scores in the recent positions with scores more like those throughout the rest of the list (with a recall probability of about .25 per item) then the total number of proverbs recalled (a summation of proportions correct across serial positions) would be about 4. By this alternative analysis, this number may approximate the number of chunks that can be well-encoded into working memory at the same time.

In our recent work (Chen & Cowan, 2005; Cowan, Chen, & Rouder, 2004), we have made several attempts to refine the methods used in the past to assess the capacity limit of working memory using verbal lists. We are no longer sure that our newer methods were successful, although they did yield important clues. Cowan et al. (2004) had participants recall on each trial, in order, the items in a list of 8 one-syllable, common words. Words in a list were always presented in pairs on the computer screen during a serial recall trial. Prior to this serial recall test, though, participants had studied the pairing of words to be used in most of the lists. Each word pairing was studied 0, 1, 2, or 4 times and these words were then used as the stimuli within lists in 0-pairing, 1-pairing, 2-pairing, and 4-pairing conditions, respectively. To equalize the number of prior presentations of the words themselves, along with the pre-exposure to pairs, the words also were presented as singletons 4, 3, 2, or 0 times in the 0-, through 4-pairing conditions, respectively, for a total of four presentations of each word (in the form of a singleton or a pair member or both) before the recall test. There were also lists in the recall phase of the experiment that were composed of nonstudied words.

This design yields predictions that appear to pit the phonological loop account against the capacity-limited account. A list of 8 singletons has as much speech material as a list of 4 learned pairs, but the latter has fewer chunks, with the number of chunks decreasing as the increasing number of pair exposures leads to a larger average size of chunks. In scoring the results, we allowed two types of chunks: (1) two-word chunks, comprising pairs of words that were presented together within a list and were recalled with the two words in immediate succession in the presented order, and (2) one-word chunks, comprising any other word recalled. A mathematical model took into account the possibility that two words could look like a single, two-word chunk in recall but actually could have been recalled as two separate one-word chunks; but that rarely happened according to the model. (The expected rate of that, based on such factors as the combination of first-word-only and second-word-only recall rates, was very low.)

The key results of this study can be observed in the top part of Table I. In contrast to the phonological loop hypothesis, the number of words recalled increased markedly with the number of word pairings. (Note that a result more consistent with a phonological loop hypothesis was obtained if the results were scored strictly according to serial position, as in Baddeley et al., 1975.) In keeping with a chunk capacity limit hypothesis, the total

TABLE I

Words Recalled and One Estimate of Chunks Recalled in Free and Serial Recall

Experiment	Condition	Words per list	Words in free recall	Words in serial recall	1- and 2-item chunks in free recall	1- and 2-item chunks in serial recall
Cowan et al. (2004)	Nonstudied (=8n)	8	–	**4.64**	–	**2.83**
Cowan et al. (2004)	0-paired (=8s)	8	–	**4.88**	–	**3.33**
Cowan et al. (2004)	1-paired	8	–	**5.50**	–	**3.38**
Cowan et al. (2004)	2-paired	8	–	**5.69**	–	**3.48**
Cowan et al. (2004)	4-paired	8	–	**6.54**	–	**3.50**
Chen and Cowan (2005)	4n	4	3.45	3.75	1.88	1.97
Chen and Cowan (2005)	4s	4	3.73	3.88	2.03	2.09
Chen and Cowan (2005)	6n	6	4.36	4.34	2.67	2.72
Chen and Cowan (2005)	6s	6	4.52	4.00	2.79	2.56
Chen and Cowan (2005)	8n	8	4.76	**4.56**	3.12	**3.16**
Chen and Cowan (2005)	8s	8	4.76	**5.00**	3.18	**3.06**
Chen and Cowan (2005)	4p	8	6.67	**6.34**	4.15	**3.28**
Chen and Cowan (2005)	12n	12	4.70	3.88	3.03	2.47

(continued)

TABLE I *(continued)*

Experiment	Condition	Words per list	Words in free recall	Words in serial recall	1- and 2-item chunks in free recall	1- and 2-item chunks in serial recall
Chen and Cowan (2005)	12s	12	5.45	3.69	3.82	2.44
Chen and Cowan (2005)	6p	12	8.73	7.63	5.21	4.06

Note. The data are drawn from Cowan et al. (2004: Experiment 1) and Chen and Cowan (2005: Experiment 1, free recall and Experiment 2, serial recall). Cowan et al. (2004) conditions: where x is 0, 1, 2, or 4, "x-paired" refers to items paired x times and shown (4 − x) times as singletons. Chen and Cowan (2005) conditions: where x is 4, 6, 8, or 12, "xs" refers to x well-learned singletons; xn refers to x nonstudied singletons; xp refers to x well-learned pairs.

number of chunks recalled (1- plus 2-word chunks) was fairly constant at about three and a half chunks across pairing conditions, with slightly fewer chunks recalled in the nonstudied word condition.

Our subsequent work (Chen & Cowan, 2005) shows that there is a limit to this observation of a fixed limit in chunks recalled; it does not hold in just the same way across list lengths. However, it also suggests that we are making progress toward understanding the role of chunking and chunk capacity limits in immediate recall. Chen and Cowan used both free and serial recall, and used list lengths of 4, 6, 8, and 12 items. The main reason to manipulate list length was to examine the effects of list length with the number of chunks varied, and to examine the effects of the number of chunks with the list length varied. In the pretraining period, each word was presented only as a singleton or in a pair. The training continued until the participant was 100% correct in cued recall on the pairs (given the first member of the pair, correct recall of the second member) and 100% correct on the singletons (given the word, correct recall indicating that it had no paired associate). That way, all trained pairs could be assumed to be learned chunks. In 4s, 6s, 8s, and 12s conditions, the words had been studied as singletons whereas in 4n, 6n, 8n, and 12n conditions, the words were nonstudied. The digit refers to the number of words (presumed one-word chunks) in the list. Critically, there also were 4p and 6p conditions consisting of 4 and 6 well-learned pairs (two-word chunks), respectively.

The metric of performance used by Chen and Cowan was proportion of words correctly recalled. Several comparisons were critical. If recall depended only on the length of the list, as the exclusive use of a phonological loop mechanism would imply, then recall should be identical for the 8s, 8n, and 4p conditions inasmuch as all of them include 8 words per list. In contrast, if recall depended on the exclusive use of a chunk-capacity-limited mechanism, then recall should be identical for the 4s and 4p conditions, inasmuch as both of them include 4 known, familiar chunks. Similarly, a length mechanism implies equality of the 12s, 12n, and 6p conditions whereas a chunk-capacity-limited mechanism implies equality of the 6s and 6p conditions.

Before introducing the results, it should be mentioned that there are different possible limitations of using free and serial recall to test capacity predictions. The traditional interpretation of these procedures is rather inconsistent. In serial recall, the traditional assumption has been that all recalled items are in working memory (e.g., Baddeley, 1986). Only short lists are typically used (e.g., 9 or fewer items) inasmuch as not many items can ever be recalled in the correct serial order. In free recall, the traditional assumption has been that only items at the end of the list are in short-term or working memory, whereas earlier list items have been memorized and occupy long-term memory (Davelaar et al., 2005; Glanzer & Cunitz, 1966). An alternative interpretation of long-term recall, however, is that some cue to recall of each list item must be saved in working memory, with the possible exception of the most recent list items. Often, the last item or items are recalled first, presumably in order to recall them without any input or output interference, so it is not clear if they have to be encoded into working memory in the same way as other items. If so, free recall might be judged a better measure of chunks held in working memory. In serial recall, participants might not be able to do their best because they are held responsible for the serial order of items whereas, as suggested above, the best strategy to recall all items that one knows may not be to recall them in the presented order, but to present the most recent items first.

In one important way, the results closely matched the chunk-capacity-limited prediction. As one can see in the upper, right-hand panel of Fig. 1, the proportion correct in the 6p condition was very close to the proportion correct for the 6s condition (and the 6n condition, for that matter), and much higher than the 12s or 12n conditions. This is very different from the phonological loop prediction and in concordance with the chunk-capacity-limit prediction. Comparing panels of the figure, one can see that the same applies to serial recall of these same lists, provided that the results are scored without regard for serial order. For other conditions, the results are not as clear and, if one examines the shorter, 4p condition for serial recall strictly scored, it matches the 8s condition as one would expect according to a phonological loop hypothesis (Fig. 1, bottom left). The conclusion appears to be that the phonological loop mechanism and a chunk-capacity-limited mechanism each have a separate role in recall and may often work together, in a balance that depends on the stimulus conditions and the data considered. When the materials fit within a period of about 2 s (as do 8-word lists, but not

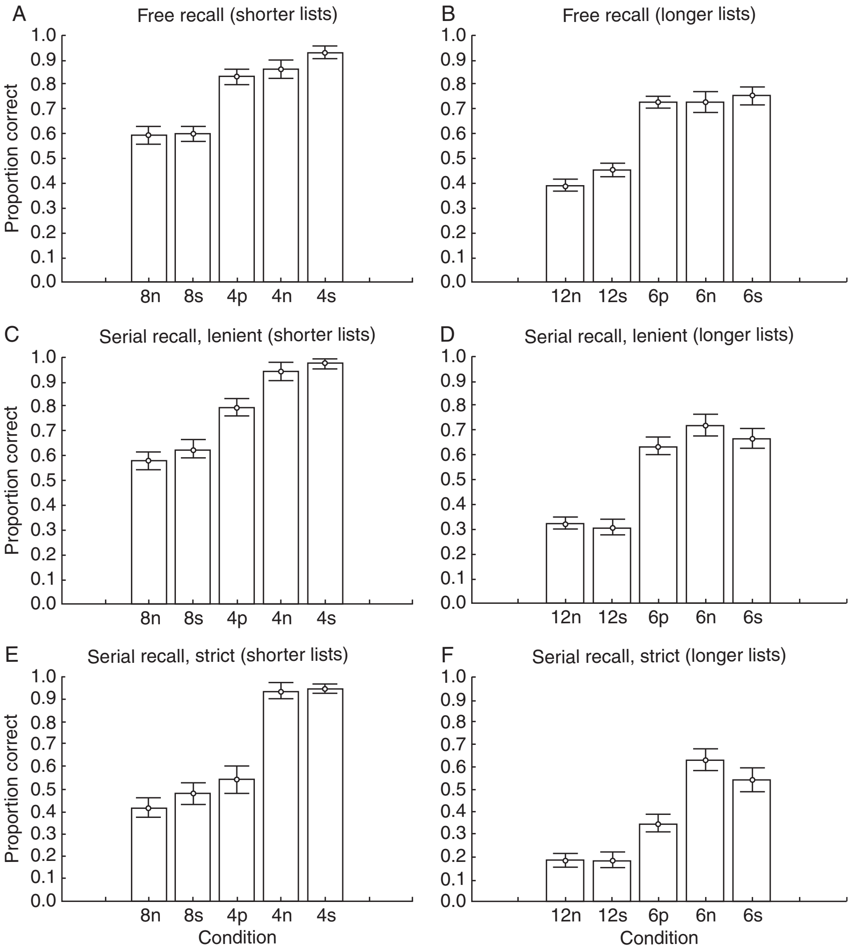

Fig. 1. From Chen and Cowan (2005, Fig. 1). Proportion correct in various conditions of two experiments. As listed in the figure, panels refer to free recall and to serial recall with strict or lenient scoring. Condition 8n = lists of 8 nonstudied words, 8s = lists of 8 words studied as singletons, 4p = lists of 8 words studied as 4 well-learned pairs, and so on. Error bars are standard errors. Capacity-limited performance is best exemplified in Panel B (free recall of longer lists); time-limited performance is best exemplified in Panel E (serial recall of shorter lists, strictly scored).

12-word lists) and when the serial order of responding is considered, it is the length of the list that counts. When the materials are longer and the scoring method ignores serial order, it is the number of learned pairs that counts. In-between conditions yield in-between results.

Given these findings, what are we to make of the results of Cowan et al. (2004), who examined the serial recall of 8-word lists and found a constant capacity in chunks? Those results were obtained with a scoring method that attempted to determine how many chunks were included in recall, regardless of the serial order of those chunks. However, there are limitations to both of the studies we have just reported. The study of Cowan et al. (2004) carefully examined chunks, but only for lists 8 words long. It is possible that the list length determines the amount of interference and therefore influences how many chunks can be reported. Chen and Cowan (2005) did vary list length, but the near-equivalence they found for the 6p, 6n, and 6s conditions was obtained with a straight proportion correct measure. We assumed that the equivalence in proportion correct meant that the same number of chunks was recalled in these conditions but that is not necessarily the case.

To provide a bridge between these two findings, we reanalyzed the results of Chen and Cowan (2005) in terms of chunks, counting singletons and intact pairs as 1- and 2-item chunks, respectively. As shown in Table I, the numbers of chunks recalled from 8-item lists in serial recall were fairly comparable for the data from Cowan et al. (2004) and from Chen and Cowan, as indicated by the bold entries. However, when the list length was changed or free recall was used, the equivalence was not as close (rightmost two columns of the table). Of course, the estimate of chunks recalled is necessarily limited by ceiling effects for 4-item lists, but the equivalence was not very close in 6-, 8-, and 12-item lists. That is true for free recall, for which the number of chunks recalled was an increasing function of list length. We also tried the scoring developed by Tulving and Patkau (1962), described above. Recall that, in this scoring method, any items recalled in a row were counted as a single adopted chunk. Again the number of chunks increased with list length. It showed the effect of learned pairing to be an increase in the number of chunks recalled as well as an increase in the average chunk size.

From this work, the question remains as to why we have not obtained a constant capacity estimate across list lengths. One possibility is that the process of verbal rehearsal is used to supplement the capacity-limited region of working memory, either by allowing memorization or by keeping some items in an active state in memory. In a still-unpublished follow-up study, we (Chen and Cowan) repeated the serial recall experiment of Chen and Cowan (2005, Experiment 2), but with the requirement that participants engage in articulatory suppression during the recall test. They were to repeat the word "the" at a rapid rate during the reception of the list. The results do show more equivalence across list lengths. Regardless of the list length, participants recalled about 3 words from lists of singletons and about 6 words from lists of learned pairs. This is consistent with the proposal of Broadbent (1975) that a basic capacity is about 3 items (or 3 chunks), and that one can observe this only in situations in which other mnemonic processes are curtailed.

C. FURTHER EVIDENCE OF VISUAL–SPATIAL CHUNK CAPACITY LIMITS

Most of the recent work on visual–spatial working memory has been conducted with arrays of objects to be retained in memory, inspired by Luck and Vogel (1997). In this task, a target array of simple objects is briefly presented on the screen (e.g., small squares varying in color). This brief presentation is followed by a probe stimulus (the same array, sometimes with a circle cue around one item; in later studies, a one-item probe). The question is whether there has been a change between the target array and the probe array, such as a change in color in one item. When one item is encircled in the probe array or a single-item probe is presented, the typical procedure is that only that item may have changed from the item in the same location within the target array. Adults can carry out this sort of task very well when there are four or fewer items in the array, and performance falls off as a function of the number of items in the array. Luck and Vogel found that performance levels were similar even when participants were responsible for four different features of bar objects (color, orientation, size, and presence or absence of a black segment in the center or "gap"). Thus, visual working memory seems to be limited in the number of objects that can be retained. This method of Luck and Vogel (1997) has an advantage over the earlier, classic method of Sperling (1960) that the items are nonverbal so that a purely visual–spatial store is probed. Often, articulatory suppression is used during the task to avoid verbal encoding of the items. However, there is little effect of articulatory suppression in this type of task (Morey & Cowan, 2004, 2005).

The ability to study capacity depends on a measure to convert the proportion correct to capacity. Obtaining such a measure may be easier for recognition procedures than it is for recall procedures because it is easier to take into account the role of guessing. (Therefore, it is unfortunate that recognition procedures have generally not been used to examine working memory capacity for verbal materials.) Pashler (1988) developed a method that logically seems to fit the method in which the probe is the whole array, in which any one item might have changed from the target array. The model assumes that for a target array with N objects, k objects are placed within working memory and can be compared with the probe array. Therefore, if an object has changed from the target array, the probability that the change will be noticed is k/N. If no change is detected, the participant guesses that it has changed with guessing rate g, which is taken as the proportion of incorrect responses on no-change trials. These assumptions lead to the formula, $k = N$ (hits − false alarms)/(1 − false alarms).

Cowan (2001; see also Cowan et al., 2005, appendix) modified this formula to be more suitable for procedures in which one item within the probe stimulus is cued or a single-item probe is presented. In these circumstances, it is assumed that the probability of knowing whether or not the probe has changed in comparison to the corresponding item in the target array is k/N, and the probability of guessing that it has changed is g. These assumptions differ from those of Pashler (1988) in that performance on no-change trials is not just based on guessing. These assumptions lead to the formula, $k = N$ (hits − false alarms). In many studies, though not all, it has been shown that the value of k increases with set size until reaching a rather constant, asymptotic level after about 4 items (e.g., Cowan et al., 2005). With such results, k can be taken to estimate the number of items in working memory and the asymptotic value can be taken to estimate the mean capacity across participants. The results typically show that between 3 and 4 items are held in working memory.

The assumption has been that each item presented in a brief visual array remains a single-item chunk. Although there is some evidence that chunking of visual items is indeed possible (Gobet et al., 2001; Jiang, Olson, & Chun, 2000), the pattern of results in which an asymptotic value of k is reached suggests that chunking is typically not an important factor with briefly presented, haphazard arrays of items.

Another issue is whether capacity is constant even for complex objects. Alvarez and Cavanagh (2004) found that the capacity estimate decreased as the complexity increased. The least complex was simple colored squares, with complexity increasing for line drawings, Chinese characters, irregular polygons, and the most complex set, cubes in different spatial orientations. However, Awh, Barton, and Vogel (2007) found that just as many items were held in working memory regardless of the complexity; it was just the perceptual resolution of the items in working memory that decreased with complexity. Thus, in a mixed target array that included some complex objects and some simpler objects, the ability to detect a change of a complex object to an exemplar of a different category (e.g., a change from a cube to a Chinese character) was no worse than the ability to detect a change from one color to another.

In a related neuroimaging study, Xu and Chun (2006) found slightly different brain areas to be involved in visual working memory, including an area that displayed activity in proportion to the number of objects in the display up to 4 (the inferior intraparietal sulcus), and other areas that displayed activity commensurate with Cowan's (2001) k measure, declining as the complexity of items in the array increased (the superior intraparietal sulcus and the lateral occipital complex). These studies show that the topic of working memory capacity is an exciting one in which the influences between behavioral and neurological sources of evidence may truly be bidirectional.

Additional work has gone into determining whether attention is needed to retain the binding between features within a visual item. One might expect so on the basis of the work on attention showing that it is much more difficult to search for a conjunction of features (e.g., a red square among red circles and blue squares) than it is to search for a single feature (e.g., a red square among blue squares only). However, a discrepant expectation could be formulated on the basis of the finding of Luck and Vogel (1997) that participants can retain just as many visual objects no matter whether the participants are responsible for one or four features of the objects. The findings seem to support the latter expectation. Studies using the Luck and Vogel procedure with simple objects have examined the effects of dual tasks on memory for features (e.g., whether the probe array contained a color that was not in the target array) and for feature binding (e.g., whether a particular color from the target array appeared at a particular location in that array). Although dual-task decrements have been observed, they have been shown to be of similar magnitude for feature memory and for feature binding memory (Allen, Baddeley, & Hitch, 2006; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006). One possible resolution of all of the results is that attention is needed to enter objects into the attention-dependent part of working memory but that this process of entering objects into this part of working memory already includes the bindings between features within each item, so that no additional exertion is needed to retain the bindings. Location may play a special role in allowing features of an object to be bound (Treisman & Zhang, 2006). If features are bound as objects enter working memory, this concept assigns a high importance to the attention-related part of working memory (cf. Cowan, 1988, 1995, 2001; Cowan et al., 2005). That theme is to be reinforced in the next section.

D. Evidence of Cross-Modal Chunk Capacity Limits

One important question is why a visual and verbal working memory capacity limits exist. Given that the limits are similar in both modalities (and in other modalities: see Cowan, 2001), a key account states that the limits stem from the use of attention as a holding device regardless of the modality. This seems simpler than the alternative possibility that there are capacity limits that apply separately in the cases of verbal and visual information or, indeed, a separate capacity limit for each feature map (Wheeler & Treisman, 2002). A simpler alternative is that there are passively held, temporary representations that are not capacity limited, such as sensory memory representations in each sensory modality (for a review, see Cowan, 1995) or phonological and visuospatial buffer representations (Baddeley, 1986). A single, central capacity-limited mechanism such as the focus of attention would be able to retrieve information from a particular type of passively held representation if it still existed at the time of recall and if it proved to be relevant to the question asked about the memoranda. The central capacity limit would define how many items could be retrieved from any such representation into the central store.

If visual and verbal information both share a central storage area such as the focus of attention, then it should be possible to show that the need to retain stimuli in both modalities results in a decrease in performance in at least one modality, compared to a situation in which only retention in that one modality is required. Saults and Cowan (2007) were able to demonstrate this. The procedures that they used in five experiments are reproduced in Fig. 2. In Experiment 1, a visual array was presented at the same time as a simultaneous acoustic array of digits spoken from four different loudspeakers arranged around the participant at ear level, in four different voices to aid in perception (adult male, adult female, child male, and child female). The use of simultaneous sounds was intended to minimize participants' ability to rehearse the sounds. After a short retention interval, an array was presented in one modality or another and the task was to determine whether there was a change in a stimulus within that modality (i.e., in the digit identity of a sound or the color of a small square). In unimodal trial blocks, participants knew in advance which modality would be tested on each trial; that was not the case in bimodal trial blocks.

The question was to what extent performance in each modality was impaired in the bimodal condition. To the extent that both modalities depend on a common, central storage mechanism, the attempt to use this mechanism for both modalities at once should limit the availability of storage and compromise performance. In the limit, if only the central storage mechanism were used, then the capacity of the visual and auditory trials for bimodal trial blocks, added together, should not exceed the capacity of whichever unimodal condition has the highest capacity. (Given that the presentation of acoustic stimuli was somewhat difficult to perceive despite our best efforts, the measured capacity in that modality was lower than in the visual case.)

The results are shown in the left-hand panel of Fig. 3. Auditory capacities are stacked on top of visual ones in this figure. The modality-specific capacities were smaller in the bimodal condition than in the unimodal condition, suggesting that there was a central resource shared between the modalities. However, the fact that the sum of bimodal capacities is still higher than the visual unimodal capacity indicates that modality-specific sources of memory also contributed, in addition to a central store. The results were very similar in Experiment 2, in which the probe stimulus always included both modalities (as shown in the second panel of Fig. 3).

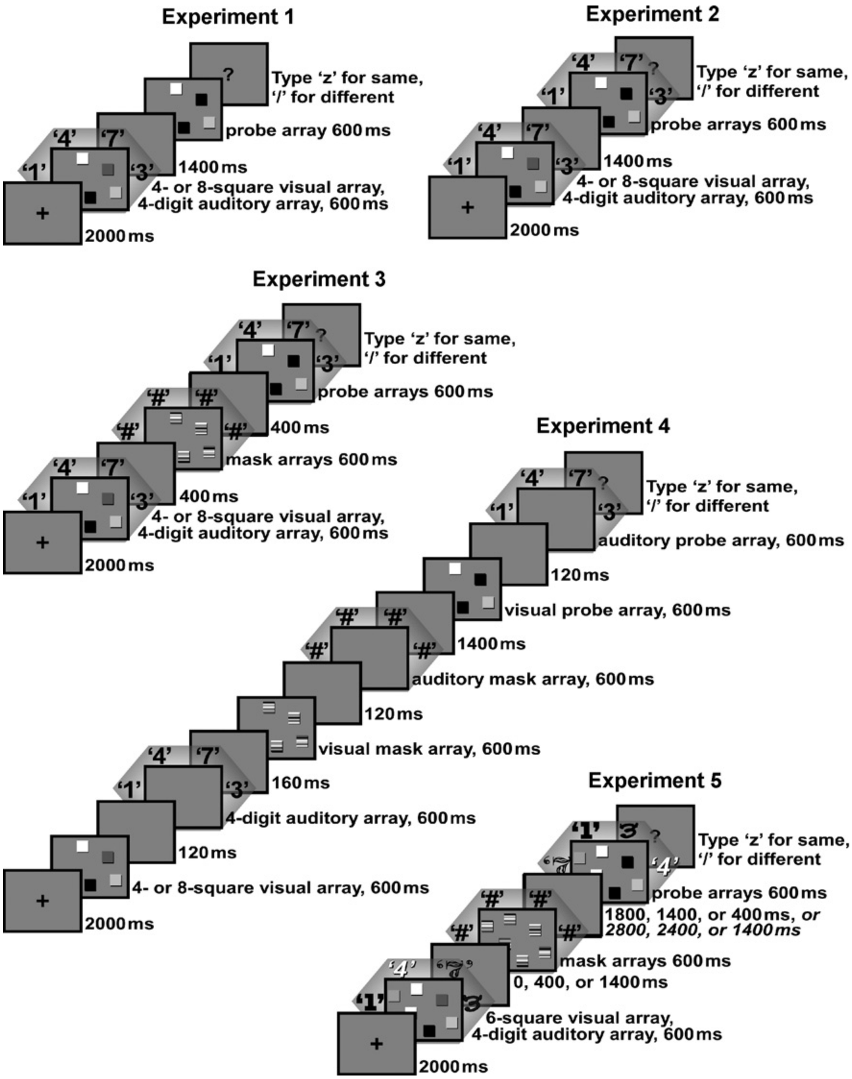

Fig. 2. From Saults and Cowan (2007, Fig. 1). Method in five experiments. Characters outside of the rectangles represent spoken stimuli, which differed in voice as represented by the different typefaces. For Experiment 4, only one of two orders of the memory sets is shown. Unimodal trial blocks were the same as bimodal trial blocks except that the participant knew which modality would be probed.

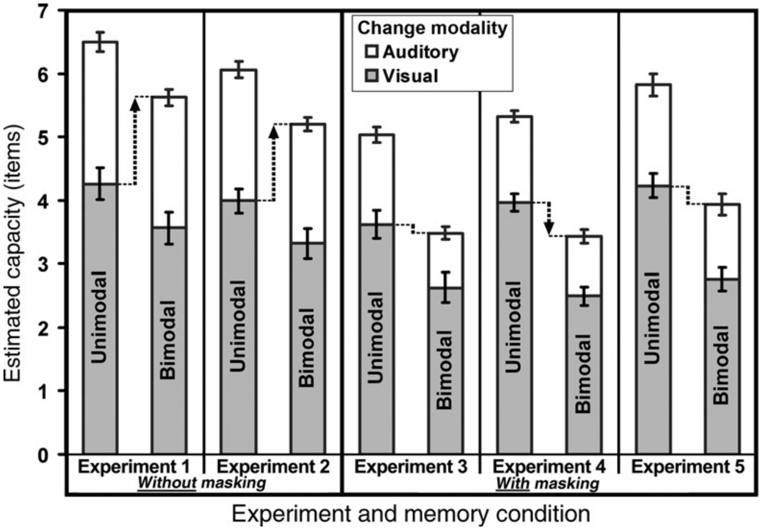

Fig. 3. Based on Saults and Cowan (2007). Results from five experiments. Provided that a postperceptual mask followed target stimuli (which occurred in Experiments 3–5), the sum of visual and auditory capacities in the bimodal condition did not exceed the visual unimodal capacity. Error bars are standard errors.

It occurred to us that we had not ruled out the possible contribution of modality-specific memory to performance. In previous studies in the visual modality, our assumption had been that the probe array would eliminate such contributions by overwriting the modality-specific representation of the target items before a decision could be made. However, that assumption had never been tested. Moreover, it might apply less to the auditory modality than to the visual modality. Perhaps participants can use the contents of capacity-limited memory and, if a discrepancy between it and the probe display is not found, they can refill the capacity-limited store with additional items from sensory memory and continue to make the comparison with the probe display. To address this possibility, in our subsequent experiments, we waited long enough after the target arrays for perception to be complete and then presented a bimodal mask on every trial, to wipe out any modality-specific memories that remained by that time.

As shown in Fig. 2, this procedure was used in Experiments 3 through 5. The mask included multicolored squares in the same locations as the squares in the visual target array and mixtures of all of the digits in each particular voice, presented through the same speakers as the target items had been presented. Experiment 3 resembled Experiment 2, but with the addition of the mask. Experiment 4 ruled out the contribution of resource limitations during stimulus encoding by presenting the target arrays in the two modalities in succession rather than simultaneously. Finally, Experiment 5 showed that the resource limitation was not simply the use of spatial codes for both modalities. In this last experiment, the spatial locations of spoken digits were rearranged between the target and probe arrays so that the participant had to rely on the association between digits and voices to respond. Moreover, different masking delays were tested to ensure that perception was complete by the time the mask was presented.

The results of three experiments using this postperceptual mask are shown in Fig. 3 (in the three right-most panels). In each such experiment, unlike Experiments 1 and 2, the sum of visual and auditory capacities in the bimodal condition did not exceed the unimodal visual (or the unimodal auditory) capacity. The close match between the unimodal visual and bimodal total capacities suggests to us that we have managed to isolate and estimate the contribution of a central capacity to working memory. On average it contributes between 3 and 4 items, in keeping with the theoretical estimates of Broadbent (1975) and Cowan (2001).

A closely related question is whether the cost of retaining two sets of materials presented in the same modality is any more costly than retaining two sets of materials presented in different modalities. In principle, the central store should not impose additional costs for materials all presented in one modality. However, the same cannot be said for the perceptual and conceptual processes needed to load materials into working memory. Cowan and Morey (2007) found a way to isolate a central maintenance component that is modality-free in this way, even though there are more modality-specific perceptual and/or conceptual processes. They included some trials with only one stimulus set and other trials with two stimulus sets. For the latter, as shown in Fig. 4, either set could be a verbal sequence or a visual array. (Articulatory suppression was used to prevent rehearsal.) After that, there was a cue with two boxes, one on top of the other, indicating that the first set should be retained (question mark in the top box), the second set should be retained (question mark in the bottom box), or both sets should be retained (question marks in both boxes). A 3-s retention interval followed and then a probe was presented for just one stimulus set, always one that the participant has been asked to retain. The results of this experiment are shown in Fig. 5. There was a cost of receiving two stimulus sets instead of just one, and that cost was larger when both sets were in the same modality. However, when two sets were presented, there was an additional cost of being asked to retain both sets rather than just one (“both cued”), and that additional cost was about the same no matter whether the two sets were in the same modality or not. Thus, after the items all had been loaded into working memory, the

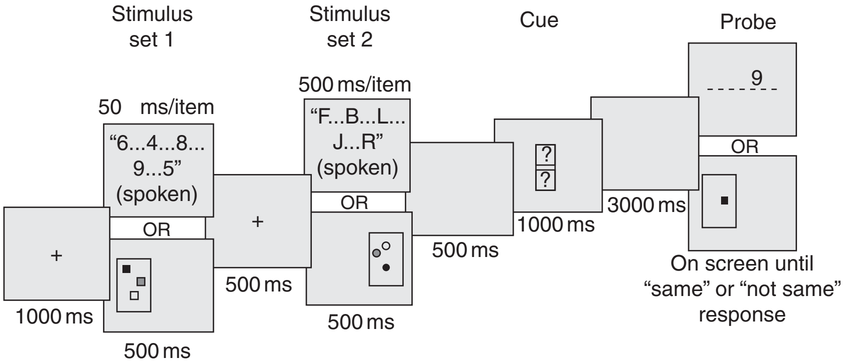

Fig. 4. From Cowan and Morey (2007, Fig. 1). Experimental procedure to isolate dual-task conflicts occurring during dual retention, as opposed to during encoding or responding. Shading of small squares or circles reflects variation in color. A spoken or visual first stimulus set is combined with a spoken or visual second set. If both sets are spoken they differ in voice and item category; if both are visual, they differ in the side of the screen and shape of objects. The postcue can indicate the need to retain the first set (top question mark), second set (bottom question mark), or both sets as shown here (two question marks, dual-retention trials). Only the processes after the cue differ between single and dual retention trials. There also were one-stimulus trials (not shown).

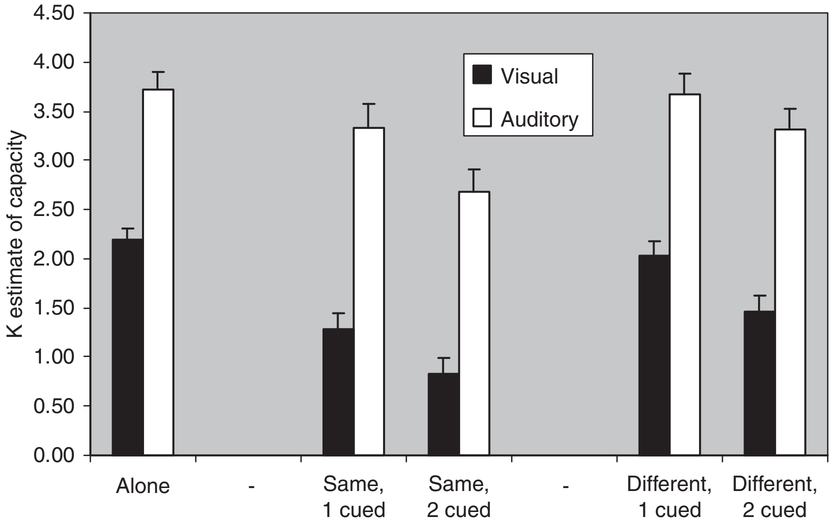

Fig. 5. Based on Cowan and Morey (2007). Visual and auditory capacities using the capacity estimation method of Cowan (2001) and Cowan et al. (2005). Error bars are standard errors. The experiment included articulatory suppression to prevent rehearsal of verbal lists. *Same* and *different* refer to whether the modalities of the two sets differ. Performance was higher when only one stimulus set was presented (*alone* condition) than when two same-modality sets were presented. However, the cost of dual retention (2 cued vs 1 cued) was about the same no matter whether the stimulus sets were in the same or different modalities.

continued retention of those items in working memory occurred in a central store that did not depend on modality.

III. Energy Limits, Time Limits, and Combinations of Limits

A. Multiple Limits

Although our own recent work has focused on space or chunk capacity limits, we do not wish to imply that these limits necessarily dominate working memory studies. Other major strands of research have produced findings more aptly described with the notion of energy or resource limits, and possibly with time limits. Bear in mind that much of the relevant evidence has focused on ideographic questions and will be deferred to the next section.

B. Energy Limits

Energy or resource limits refer here to processing limits that hold across many types of attention-demanding activities and must be shared between these activities. Often, it is a matter of judgment whether to classify a result as capacity limited or resource limited. The reasons are that (1) there appears to be a trade-off between capacity and resources used for other purposes, and (2) we often have insufficient evidence to determine which fundamental limit is restricting performance. Regarding the former (a trade-off between capacity and other resources), work in several laboratories confirms the use of attention at least in the case of visual working memory. Morey and Cowan (2004) used the procedure illustrated in Fig. 6, and found that an array-comparison task was impeded by the need to recite a random seven-digit load during the period between arrays, but was not affected by the need to recite a known seven-digit telephone number during that period. The effect of a seven-digit load was especially large when an error was made on the digit load. Moreover, the effect of the seven-digit load depended on that load being repeated aloud rather than silently rehearsed (Morey & Cowan, 2005). Stevanovski and Jolicoeur (2007) found that a simple tone identification task can impede array comparisons. These results suggest that attention must be shared between visual working memory and various sorts of retrieval. When little or no interference is found between a visual load and a verbal load (e.g., Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002), that may be because it is possible to use a separate verbal rehearsal process to maintain the verbal information silently, and the rehearsal process is thought to require relatively little attention in young adults (Guttentag, 1984). To observe interference between tasks, one must not be able to rely upon an automatic, effortless process in either task.

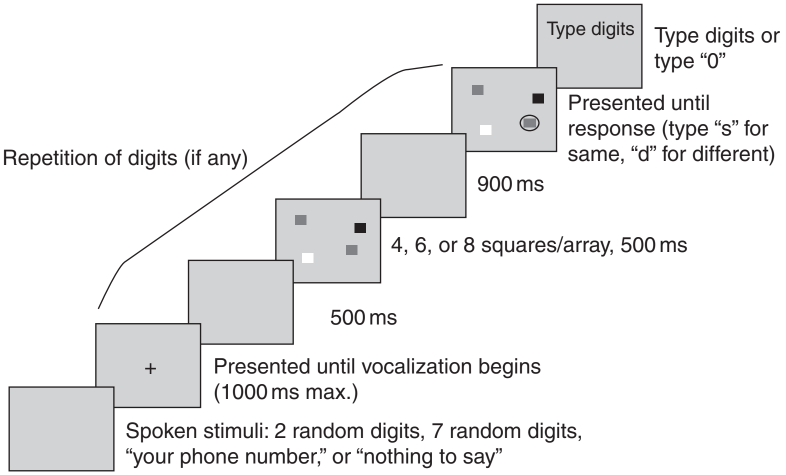

Fig. 6. From Morey and Cowan (2004, Fig. 1). Procedure to examine effects of a memory load, and of articulatory suppression, on array comparison performance. There was an effect of a random 7-digit memory load but not of a known 7-digit number (the equivalent suppression condition).

The converse trade-off effect may also be the case; there is an effect of a working memory load on nonmemorial tasks that rely on attention. Lavie, Hirst, de Fockert, and Viding (2004) introduced a theory that attention depends on the working memory load. They used an attention task in which the participant was to identify a target letter as x or z. It was accompanied by a distractor letter that was compatible (e.g., X with target x), incompatible (e.g., Z with target x), or neutral (e.g., L with target x). Meanwhile, one or six letters were held in memory for a search task. It was found that the incompatibility effect was larger with a high memory load. The interpretation was that a working memory load uses up resources that otherwise would be used to inhibit the irrelevant distractors.

More recent work has suggested that the effect is not consistent and that a working memory load can reduce, as well as increase, distractor interference (Kim, Kim, & Chun, 2005). As these authors pointed out, working memory is not unitary so it is possible to design situations in which the distractor has more in common with the working memory load (such as verbal processing) than the target does.

From this past research, it is clear that we need to know the results of a procedure in which the working memory task is limited to its attention-related components. It should then only increase distractor effects, as in Lavie et al. (2004), instead of sometimes decreasing distractor effects as in Kim et al. (2005). Even if that prediction is confirmed, it will not be clear how to interpret the attention-related working memory component(s). It may be, as we have suggested above, that attention can be used directly to hold some information in its focus. An alternative possibility more in line with the classic model of Baddeley (1986) is that central executive processes are needed to keep the items refreshed in more passive stores. In terms of Baddeley (2000), the question is whether attention-related storage occurs in an episodic buffer or occurs in more passive buffers with the assistance of central executive processes. So, more research is needed to understand the attention-related part of working memory.

A second reason why it is still a matter of judgment whether to classify a result as capacity limited or resource limited is that we have insufficient evidence on the nature of limits. Consider a classic procedure that has been used to examine resource limits, the procedure of Stroop (1935) in which color words are presented in different colors of ink and the task is to ignore the word and name the color of ink. Individuals are much slower and more error-prone in this task compared to a task in which the ink is presented in the form of noncolor words or nonwords. A capacity-based account of this finding is that, on some trials, the participant has dropped the goal from working memory and either erroneously makes the more natural, prepotent response of reading the word, or requires time to reload the goal into working memory before responding. A different, resource-based account is that the goal remains in working memory but nevertheless requires attention-related resources to overcome the difficulty of resisting the prepotent response, and to inhibit it while carrying through with the relevant response. Many results that are typically thought of as resource related thus can be alternatively described as capacity-related. To do so, however, one cannot simply think of capacity in equal chunks. If the only temporary memorandum of the Stroop task is the task goal, then this goal would have to count as more than one chunk in order to provide consistency with the literature above, indicating that several chunks can be held in working memory at once. Once one gets into the concept of a variable size of chunks in working memory, what is needed is a fluid resource that can be apportioned among the pieces of information as required. Attention would appear to be such a resource.

C. TIME LIMITS

The field of working memory seriously shifted from an interest in capacity limits to an interest in time limits when Baddeley et al. (1975) published their work on the word length effect. Lists of short words were recalled better than lists of long words even when the short and long words were equated for the number of phonemes and syllables. Individuals recalled about as many verbal items as they could recite in about 2 s, so it was suggested that working memory is dependent on a decaying store that becomes useless within 2 s unless it is refreshed by a rehearsal process. The faster it can be refreshed, the more items can be held in storage at the same time and then recalled when the list ends.

Currently, however, there is no general agreement that the word length effect is obtained across different sets of materials when the short and long materials are equated for the number of phonemes and syllables (see Mueller, Seymour, Kieras, & Meyer, 2003; Neath, Bireta, & Surprenant, 2003) so there is no agreement that the word length effect supports the notion of time limits per se.

Some recent work using complex spans that include storage and processing components (operation span and reading span) have shown that time is an important variable for forgetting (e.g., Towse, Hitch, Hamilton, Peacock, & Hutton, 2005). Nevertheless, this is time filled with distracting items in order to prevent rehearsal. Given that these can cause interference, it is not clear that temporal decay is the cause of forgetting. Indeed, Saito and Miyake (2004) found that it is the number of distracting items rather than the time that is critical for performance.

Barrouillet, Bernardin, and Camos (2004) more specifically showed that it is the density of distracting events that is critical (i.e., the number of retrievals divided by the time in which those retrievals have to take place). They inserted numbers to be read aloud between items (letters) to be recalled. The faster these distracting numbers were presented, the poorer recall was. Their theoretical explanation involves a decaying memory that can be rehearsed during the breaks between distracting items. However, it is noteworthy that they do not find a large effect of time per se, as opposed to just the density of distracting events. If a certain density of distracting events is enough to prevent full rehearsal of the stimuli, there should be a memory leak that should be manifest as poorer performance with more time at a certain distraction density.

More recently, Lewandowsky, Duncan, and Brown (2004) carried out more direct tests of time limits. The amount of time between one item and the next in the recall period of a serial recall task was manipulated and, despite procedures designed to prevent rehearsal, no significant effect of time was obtained. The authors suggested that information is not lost as a function of time during the response, even when it is not rehearsed. A similar result was obtained by Cowan, Elliott, et al. (2006). Second-grade children were asked to recall items more quickly than usual and succeeded in recalling items as quickly as adults ordinarily do; yet their memory was not improved at all by this manipulation. We are currently conducting follow-up research to cross-pollinate the research procedures of Lewandowsky and colleagues versus Barrouillet et al. (2004).

Cowan and AuBuchon (2008) did find an effect of time in the absence of external interference by presenting irregularly timed lists; a long first half-list hurt performance, but only if the list timing had to be reproduced in the response. The assumption was that reproducing the timing prevents rehearsal and takes up attention during the response. In such circumstances, information can drop out of working memory during reproduction of a long first half-list. However, the detrimental effect of reproducing the timing of a long as opposed to a short first half-list occurred only in restricted circumstances (in the presence of a long second half) and its basis is not yet clear. Cowan and AuBuchon did not attribute it to a decay effect per se.

In sum, it seems fair to state that there is, practically speaking, an effect of time, though it may well result from the effects of interference spread across time (e.g., in the acid bath theory of Posner & Konick, 1966; see also Massaro, 1970) rather than from a pure process of decay even in the absence of interference. In the word length effect, each phoneme or syllable could interfere to some extent with memory for previous phonemes or syllables.

D. Combinations of Limits

If two or three different types of basic limitation on working memory exist, there is no reason why a particular procedure should tap only one of these limits. There are two ways in which different types of limits might be combined in working memory. The two limits can occur because they stem from a common mechanism, or the two limits can occur because different mechanisms affect performance separately. An example of two limits possibly stemming from the same mechanism and both affecting performance is the combination of space (chunk capacity) and energy (resource) limits. Both of these could be based on a common attention-related resource. A common attentional resource might have to be split between maintenance of items in working memory and central executive functions such as shifting attention, updating working memory, and inhibiting irrelevant items (e.g., Miyake, Friedman, Emerson, Witzki, & Howerter, 2000).

An example of two limits possibly stemming from different mechanisms and affecting performance together is the combination of space (chunk capacity) and time (decay) limits. Chen and Cowan (2005) found that the length of lists determined recall for short lists with serial scoring, but that the number of chunks governed recall for long lists without serial scoring. As shown in Fig. 1, the result was intermediate for other conditions. As we mentioned, in a follow-up study in which articulatory suppression was added, the chunk-based limit determined recall much more uniformly across list lengths. All of this suggests that there is a capacity-limited mechanism that governs performance except insofar as it is possible to engage in verbal rehearsal for the serial recall of short lists, and that the two mechanisms both can be used together for the same recall procedure (see also Mulligan & Lozito, 2007).

IV. Ideographic Evidence

A. USING IDEOGRAPHIC EVIDENCE

We have already discussed (in Section I.C) the point that ideographic information is collected for several different and interlocking reasons. One is to find out about the classes of individuals being tested, but another is to shed light on nomothetic processes. The notion behind this approach is that if individuals differ in Trait X, then Trait X is nomothetically important too. An example of this line of reasoning is the idea that if individuals with high versus low working memory differ in their ability to control attention, then the control of attention is an important source of the human limitation in working memory. The paper by Kane, Bleckley, Conway, and Engle (2001) is a prime example. They used a *prosaccade* task requiring that the participant look at an appearing object, and an *antisaccade* task requiring that the participant look in the opposite direction. The antisaccade task, which is quite unnatural, is accomplished more successfully by high-span than by low-span individuals, who do not differ in the prosaccade task.

Despite the strength of such findings (for convergent evidence, see Conway, Cowan, & Bunting, 2001; Kane & Engle, 2003), several related points must be considered, as follows:

(1) Any Trait X distinguishing high- and low-span individuals may not be the only difference between the groups or individuals relevant to working memory. (Example: individuals may also differ in chunk capacity, which might be very important for the human limitations in working memory.)

(2) Trait X, though correlated with working memory capacity, may not be a key determinant of working memory. (Example: it could be chunk capacity rather than the control of attention that is critical for working memory capacity limits, if chunk capacity and the control of attention are highly correlated.)

(3) Trait X in a nonstandard population could impose a limit on working memory that is not critical for normal adults. (Example: vocabulary knowledge is not ordinarily a limiting factor for verbal working memory task performance in adults given the simplicity of vocabulary selected for the task, but it probably is a factor in the performance of young children on the same task.)

(4) Some other Trait Y may be a nomothetically important limit of working memory even if it does not yield individual differences. (Example: If it were found that chunk capacity did not differ at all among individuals, it

would still be possible that all individuals are limited to, say, 4 concurrently held objects in the central portion of working memory.)

As we will show, there are ways to begin to sort out these possibilities with experimental manipulations.

B. INDIVIDUAL DIFFERENCES AMONG YOUNG ADULTS

There are distinct individual differences on many cognitive traits and many of these are correlated with various working memory tasks. However, there has perhaps been too much of an attempt to seek evidence confirming particular accounts of the correlations and not enough seeking evidence of disconfirmation. For example, Baddeley et al. (1975) found a high correlation between the rate at which an individual could recite words and the number of words of the same type that the individual could recall. Their word length effect can be viewed as the result of an experimental manipulation affecting the rate at which a given individual could recite the words, and it did seem to confirm that rehearsal speed was important for serial verbal recall (cf. Schweickert, Guentert, & Hersberger, 1990). By equating short-word and long-word lists on the number of phonemes and syllables, there was an attempt to show that it was time per se that was important rather than a correlated difference between short and long words. (Recent research challenges the success of that control; see Neath et al., 2003.)

What was not established by Baddeley et al. (1975) and perhaps has not been established in subsequent research is the boundary conditions for the effect of individual rehearsal speed on working memory. There are many types of working memory that theoretically should not depend on rehearsal speed at all. We may not have such information at present. However, Conway, Cowan, Bunting, Therriault, and Minkoff (2002; see their Table 2) examined other kinds of simple processing speeds and the results indicate that these (in particular, for copying digits and letters, and for comparing pairs of patterns and letter strings) correlated with short-term memory in the presence of articulatory suppression much better than they correlated with short-term memory without suppression. This raises the possibility that the correspondence between individual speech rates and memory spans obtained by Baddeley et al. may have an unexpected path of causal relation.

The suggestion that the control of attention is causally related to working memory span (e.g., Kane et al., 2001) has been aided by research manipulating attention. This research shows that dividing attention during a memory retrieval task has a much larger detrimental effect on individuals with a high working memory span than it does on those with a low working memory span (Kane & Engle, 2000; Rosen & Engle, 1997). In fact, dividing attention removed most of the difference between the response patterns of high- and low-span individuals. That is exactly what would be predicted if the difference between high- and low-span individuals is that the high spans make more use of attention in the retrieval task, making them more vulnerable to divided attention.

Nevertheless, there are limits to what can be concluded from this kind of result alone. For example, it could be that dividing attention saps both space (chunk capacity) and energy (attention deployment) in working memory and then it would be unclear which is critical to task performance. By analogy, suppose that one person's home computer carried out calculations better than another's. Suppose, though, that a special program in the better computer, unavailable in the poorer computer, required an internet connection to do the best calculations. If one knocked out the internet connection in both houses and observed that the calculations were now equally mediocre in both computers located in the respective houses, one might incorrectly conclude that the difference between houses was in the ability to use the internet generally, when actually the difference is in the programs available in the two computers. In this analogy, attention is the internet and chunk capacity is the computer.

Cowan, Fristoe, Elliott, Brunner, and Saults (2006) carried out a study that included measures of both capacity (visual array comparisons) and attention. The measure of the latter was obtained in a situation in which spoken and printed stimuli were presented at the same time. The instruction preceding each trial was to attend to one of these modalities, and the measure of the use of attention was the extent to which memory for the to-be-attended modality exceeded memory for the to-be-ignored modality. These measures turned out to be correlated with one another ($r = .34$), but separate, among adults. They both correlated well with a composite IQ measure (visual array comparisons, $r = .52$; attention benefit, $r = .47$). However, the variance in IQ that they picked up was largely separate. Together they picked up 12% in shared variance in IQ, the arrays task uniquely picked up another 15%, and the attention benefit uniquely picked up another 10%, for a total of 37% of the IQ variance accounted for. In light of the importance that has been placed on accounting for variance in IQ (e.g., Engle, Tuholski, Laughlin, & Conway, 1999), this is an important indication that size as well as energy limits are worth discussing.

There has been research suggesting that the relation between working memory and intelligence (or, more narrowly specified, fluid intelligence) is increased by the presence of proactive interference from previous trials (e.g., Rosen & Engle, 1998). In both an operation span task and a probed recall task, Bunting (2006) included proactive interference, in the form of stimuli similar to the ones to be recalled appearing earlier in the list or appearing in recent lists. The correlation between memory scores and fluid intelligence (Ravens Progressive Matrices) was higher when proactive interference was present. In an fMRI study using a recognition procedure, Gray, Chabris, and Braver (2003) found that high-span individuals showed more activity than low-span individuals in the prefrontal cortex, but only on trials in which proactive interference was present in the form of a lure that was not in the current list but was in a recent list. These procedures clearly show the importance of proactive interference for individual differences in working memory but the exact theoretical interpretation of the finding remains unclear. It is not clear whether what is key is the ability to suppress memory for the items in previous trials (presumably a function of attention control), or the ability to remember whether an item was presented in the current trial (possibly a function of chunk capacity).

Similarly, the finding that low-span individuals have poorer control of attention in an antisaccade task (Kane et al., 2004) could be interpreted either as their inability to remember the unnatural task goal, or their inability to override the prepotent prosaccade response. In other studies, what is observed is the specificity with which relevant items are selected (Bleckley, Durso, Crutchfield, Engle, & Khanna, 2003) or the efficiency with which irrelevant items are excluded (Heitz & Engle, 2007). In each case, though, there remain several slightly different interpretations of the findings. It could be that the difficult goal was not maintained. This failure of maintenance could occur either because it was difficult to hold the goal while using working memory for other things during the task (a space limit), or because it takes considerable effort to maintain a goal (an energy limit). Another possibility is that the goal was still maintained in working memory but that the prepotent response (e.g., making a prosaccade) was so powerful that the goal could not control behavior. This would constitute a second type of energy limit.

A study by Delaney and Sahakyan (2007) seems to favor memory limits over energy limits. It investigated the finding that high-working-memory-span individuals do a better job of directed forgetting than do low-span individuals. One explanation is that the high spans are able to use more resources to suppress the material to be forgotten. An alternative explanation is that high spans have a richer use of context in the mnemonic encoding of stimuli and are therefore better able to separate the material to be forgotten from the material to be remembered. Evidence favoring the latter view was obtained in a word list memory study. If the participants were not told to forget the list but it was followed by a different task that greatly changed the context (45 s of imagining walking through their parents' house and describing the house and furniture to the experimenter), the change in context impaired memory of the word list more for individuals with higher operation spans and, in a replication, for individuals with higher counting spans.

Work by Saxe et al. (2007) provides conceptual support for this type of interpretation from a very different source of evidence. A type of X radiation was used in adult mice, and it had the effect of impeding neurogenesis in the hippocampus. This manipulation actually assisted memory for a radial arm maze in a particular situation, in which the memory load was low but the amount of proactive interference from previous trials was high. The notion is that, in some situations, rich contextual encoding can be an impediment to recall. The richness of contextual encoding in humans apparently is correlated with working memory, so better working memory can be viewed more as a mnemonic difference than as a resource difference. For a working memory task, this mnemonic difference may translate into a higher basic chunk capacity.

Vogel, McCollough, and Machizawa (2005) pitted space (capacity) and energy (attention control) constraints against one another in a version of the visual array comparison task designed to produce a lateralized, event-related potential signal indicating storage of items in working memory. There were arrays on the left and right sides of the screen and the task was to focus attention on one of these arrays. Within that task, there were 2 or 4 items. When there were 4, they all could be relevant or only 2 of the 4 were relevant (e.g., remember the orientations of 2 red bars and ignore 2 blue bars). If the irrelevant items were easily filtered out using attention, the 2 relevant, 2 irrelevant situation should yield a potential similar to the 2-item trials, whereas if the irrelevant items could not be filtered out and had to be held in working memory along with the relevant items, they should yield a larger potential similar to the case of 4 relevant items. It turned out that the irrelevant items were filtered out easily by high-span individuals but not by low-span individuals. However, questions remain about this result. It required selection of one half of the screen to do the task in every condition, and the split between high- and low-span individuals was based on this same task. Perhaps the results would be different if individuals were categorized on the basis of a different working-memory task that did not include an attentional filtering component in every condition. On the other hand, perhaps there really is a critical resource difference between high- and low-span individuals, in addition to a chunk capacity difference.

Overall, then, it seems that some aspect of individual differences related to attention is critically important in characterizing those who yield good versus poor performance on working memory tasks and on intelligence tests. (Intelligence and fluid intelligence are so highly correlated that it has been difficult to tell whether the contributions of working memory are specific to fluid intelligence, though see Hambrick & Engle, 2001.) The notion that a general attention component is involved is a convergent result, with generality across modalities or processing domains observed in the nomothetic evidence (e.g., Cowan & Morey, 2007; Saults & Cowan, 2007) and in a latent variable analysis showing that the relation between fluid intelligence and working memory is general across processing domains (Kane et al., 2004). It is less clear whether to interpret the individual differences as fundamentally space related, energy related, or both.

C. Childhood Development

We think it has been considered less convincing to use childhood developmental differences to discern nomothetic principles, as compared to individual differences among young adults. People generally realize that the nature of the processing limitations that young children have is, in some ways, very different from the nature of the limitations that young adults have. Nevertheless, we think there are some implications of developmental research for the nomothetic principles.

The observation that adults who rehearse more quickly recall more items in short-term memory tests was followed by the observation that, as children mature, they speak more quickly and remember more in these tasks (e.g., Hulme & Tordoff, 1989). It has not been possible to teach children to rehearse more quickly to determine whether they would recall more (Hulme & Muir, 1985). Anyway, such an attempt is confounded inasmuch as rehearsal requires more attention in younger children (Guttentag, 1984). However, it has been possible to get second-grade children to speed up verbal recall responses to a rate similar to what adults usually use, and this manipulation makes no difference for memory span (Cowan, Elliott, et al., 2006). Therefore, similar to the research with young adults, we believe that the correlation between speech rate and memory span across age groups may occur for reasons other than faster refreshing of a phonological memory trace before it can decay. We do not yet have a satisfactory account of that correlation.

Recent research has focused on the possibility of age differences in capacity in visual array comparison tasks modeled after Luck and Vogel (1997). An apparent discrepancy exists in the results. On one hand, Ross-Sheehy, Oakes, and Luck (2003) suggested that infants retained about 4 items by the time that they were 10 months old, similar to the presumed capacity in adults. This suggestion was based on a procedure in which arrays appeared on both sides of the display but changes in the arrays occurred only on one side. By 10 months, infants reliably looked more often toward the changing display when it included 4 objects, but not 6 objects. In contrast, Cowan and colleagues (Cowan et al., 2005; Cowan, Elliott, et al., 2006; Cowan, Fristoe, et al., 2006; Cowan, Naveh-Benjamin, et al., 2006) have used the standard procedure of Luck and Vogel and have found a marked increase in capacity from the early elementary school years to adulthood. How can both of these results be correct?

Assuming that the development of capacity from infancy to adulthood is monotonic (a very safe assumption, we believe), either the infant research provides an overestimate of capacity or the child research provides an underestimate. The infant research could provide an overestimate because of its necessarily atypical procedure. To detect the changing side of the display, it is not necessary that all changes be detected. It is possible to model performance in the infant procedure with a process in which capacity is relatively low (e.g., 2 items) and the proportion of changes noticed is enough to motivate sustained looking when there are 4 items in the array (in which case, half of the changes in successive arrays would be noticed when these arrays were examined), but not when there are 6 items (in which case, only a third of the items in successive arrays would be noticed when these arrays were examined). On the other hand, the child research could be an underestimate because children may not fill the available working memory as consistently as adults. We have found that the children do about as well as the adults when the set size is 2 items, so it is not simply that the children more often fail to pay attention to the stimuli. Nevertheless, it is possible that children load items into working memory more slowly and therefore fail to fill their capacity before the display ends. If so, children might do as well as adults if a much longer display time was used.

One useful research strategy is to find a way to nullify the differences between children and adults. If these differences can be nullified by eliminating certain processes in adults, then the processes that were eliminated may be the ones responsible for the developmental change. This strategy is analogous to the one that Rosen and Engle (1997) and Kane and Engle (2000) used to show that dividing attention can nullify high- versus low-span differences between young adults. Cowan, Naveh-Benjamin, et al. (2006) divided attention in an array comparison task with a tone identification task and found that dividing attention made the adults perform in a manner slightly better than third-grade children but worse than fifth-grade children. In another type of procedure, Cowan, Saults, and Morey (2006) examined memory for the associations between printed words and their locations on the screen. Several words (proper names) were presented one at a time, each disappearing from the screen shortly afterward, and then a probe word presented centrally had to be dragged to the location at which that word had appeared originally. There were several different trial types and a complex pattern of performance that was very age specific. However, the addition of articulatory suppression in adults resulted in a dramatically changed pattern of performance that closely resembled that of third-grade children (who did not receive suppression), as shown in Fig. 7. Without suppression, adults seemed to engage in a process in which the list of names and the path of spatial locations were rehearsed separately, favoring trials in which there was a one-to-one

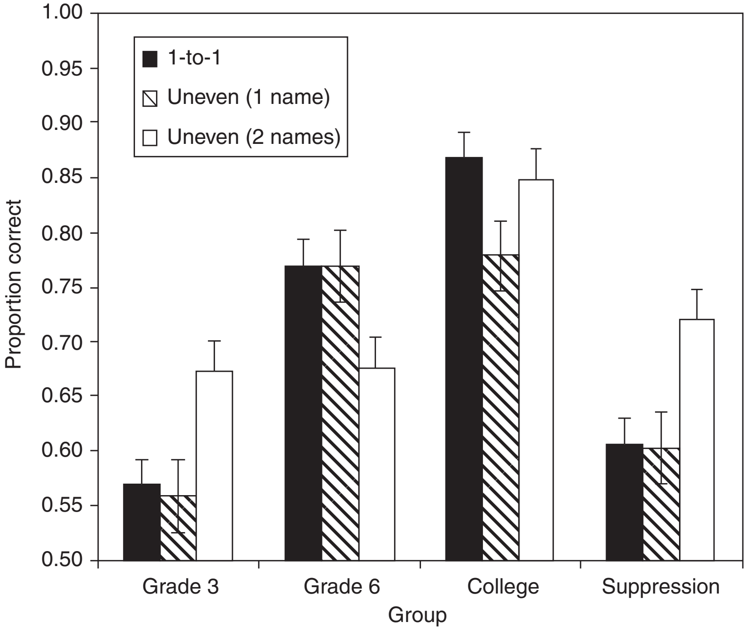

Fig. 7. From Cowan, Saults, et al. (2006, Fig. 3). Proportion correct recognition of the spatial location of a name from a target set. In the *one-to-one* condition, each name was presented at a unique location one at a time whereas, in the *uneven* condition, some locations received 2 names (not in immediate succession) on the same trial. There were 3–7 names in each target list. The two uneven conditions refer to trials in which the probed name was at a location by itself or with another name. Adults (college students) with articulatory suppression (right-hand cluster of bars) performed in a manner quite similar to third-grade children, which, as shown, was not the case for adults without suppression. Error bars are standard errors.

correspondence between the two. However, that pattern did not occur in children or in adults under articulatory suppression. Instead, they did best on trials with fewer locations used overall, even though some of those locations were used for two names on those trials. Taken together, these results suggest that strategic aspects of performance may well distinguish children from adults.

Nevertheless, there appear to be age differences in memory even when strategies are removed. Cowan, Nugent, Elliott, Ponomarev, and Saults (1999) investigated children's memory for spoken lists of digits that were presented while the children were busy playing a silent computer game that required thinking of rhymes. An occasional cue indicated that the computer game should be interrupted and the last list should be recalled. This procedure should minimize the ability to use strategies to encode the list items. In a

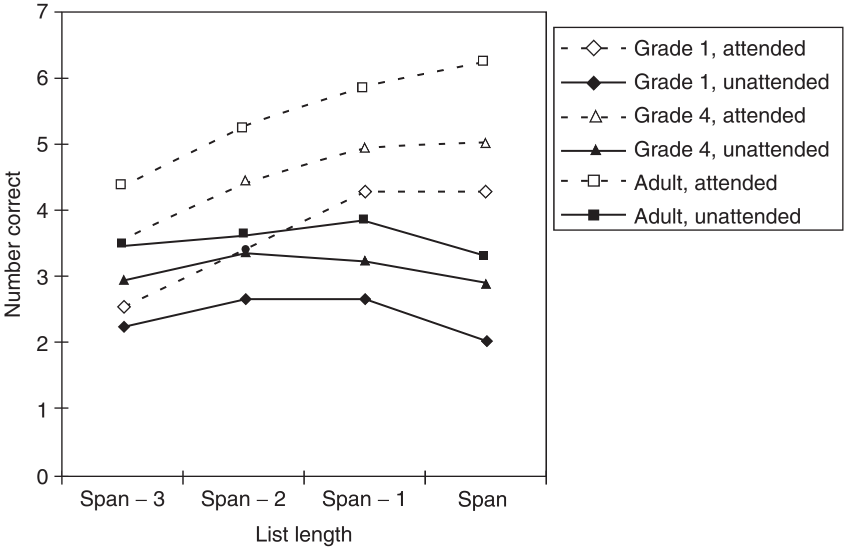

Fig. 8. From Cowan et al. (1999, Fig. 1). Number of items correct from unattended lists (solid lines) and attended lists (dashed lines) of spoken digits, in three age groups. Only lists unattended at the time of their presentation show roughly constant numbers correct across list lengths, presumably because of the absence of strategic encoding and the transfer of items to the focus of attention only after the retrieval cue.

control procedure, the lists were attended during their presentation. For unattended lists, presumably the focus of attention can be turned to the list only when the recall cue is presented. Figure 8 shows the results for attended and unattended lists. For unattended lists, the number of items recalled in the correct serial positions was constant across list lengths and increased with age. For attended lists, the number of items recalled was not constant across list lengths, and the age differences were slightly larger. However, most of the age difference in this procedure cannot easily be attributed to strategic differences during encoding. Cowan, Elliott, et al. (2005) replicated this result and obtained a similar result for a running memory span procedure, in which 12–20 digits were presented at a rapid, 4-per-second rate and ended at an unpredictable point, making rehearsal impossible. Both procedures showed similar age differences and correlated well with intelligence. For children too young to engage in strategic processing, but not for older children or adults, simple digit spans also correlated well with intelligence.

We do not yet have a reconciliation of results suggesting the great importance of strategic processing (Cowan, Saults, et al., 2006) and results showing marked capacity differences that cannot be attributed to strategic differences (e.g., Cowan, Elliott, et al., 2005). However, it can be seen in Fig. 7 that adults in the study of Cowan, Saults, and Morey with rehearsal suppressed still did recall slightly more items than young children recalled, so the contradiction may be more apparent than real.

There was an earlier era of research by neoPiagetians suggesting that children have fewer slots in working memory than adults (Case, 1972; Pascual-Leone, 1970, 2005; Weiss, 1995). Perhaps the clearest result supporting that idea is that of Burtis (1982). He presented matrices of letters to be recalled in their correct locations. The letters were presented in pairs within the matrix and the items within a pair were random or, in other conditions, contained redundancy in various ways (repetitions of the same letter; pairs presented in red to be distinct from the other pairs; pairs presented repeatedly throughout the experiment; pairs forming familiar acronyms, such as FM). The results were examined with the help of a simple model that incorporated results from both singletons and pairs. At every age for every set size, the results fell neatly on the line expected according to the capacity model with a capacity of 4 at 10 years of age, 5 at 12 years of age, and 6 at 14 years of age. The reason for the larger capacity estimates suggested by Burtis and other neoPiagetians compared to Broadbent (1975) and Cowan (2001) is an important unknown, but it might be attributed to the availability of rehearsal strategies in the procedure that Burtis used, or to other studies' use of procedures more challenging than Burtis' was, using up more capacity for processing.

In sum, there are some critically important questions remaining about the nature of the dramatic increase in working memory capabilities that are so evident between infancy and adulthood. We cannot yet be sure whether there are differences in space, energy, or both during childhood development but the evidence seems to favor both. Here we have emphasized space (capacity) differences, which are less commonly acknowledged in the literature than energy (resource) differences.

D. ADULT AGING

In some ways, adult aging may be the converse of child development. For example, whereas the rate of verbal rehearsal and verbal short-term memory both increase during childhood as noted above, in adult aging the rate of rehearsal slows down and verbal short-term memory suffers accordingly (Kynette, Kemper, Norman, & Cheung, 1990).

In other ways, adult aging may not be the converse of child development, and the differences in mechanisms of processing make it difficult to make fair comparisons across age groups in capacity. In particular, older adults appear to have more difficulty than young adults associating or binding together the features within an object (Chalfonte & Johnson, 1996; Mitchell, Johnson, Raye, Mather, & D'Esposito, 2000) or objects presented together (Naveh-Benjamin, 2000; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003). Given that one instance of binding is the formation of new chunks of information (e.g., Miller, 1956) we might expect that, in some situations, older adults will have to recall smaller chunks than young adults. To do so, they would have to recall fewer chunks even if the chunk capacity is the same in the two groups.

In several of our recent studies, we attempted to evaluate both the formation of associations and the capacity. Cowan, Naveh-Benjamin, et al. (2006) presented children (in third and fifth grade), young adults, and older adults with a visual array comparison task, similar to that used by Luck and Vogel (1997). For half of the trials in this task, the probe array did not differ from the target array (i.e., no-change trials). The probe array could differ from the target array in two possible ways for the other half of the trials (i.e., change trials). On some change trials, a colored square changed to a color that had not been seen in the original array; as a participant only had to keep track of the squares, these were considered item-change trials. On other change trials, a colored square changed to a color that was already present elsewhere in the first array. In this case, participants had to keep track of not only the colors present in the target array, but also the locations of each color. As correct detection involved successful binding of color and location features, these were termed binding trials. A cue encircled one item in the probe array, and only that item might have changed between arrays. Some key results are shown in Figs. 9 and 10. Figure 9 depicts results of an

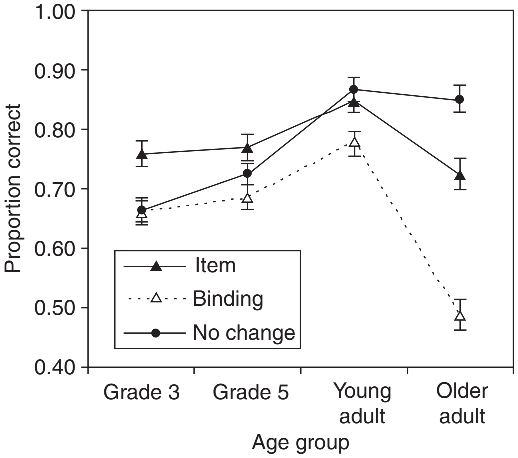

Fig. 9. From Cowan, Naveh-Benjamin, et al. (2006, Fig. 3). Proportion of trials correct for each age group (*x*-axis) in each condition (graph parameter) within Experiment 1a. Item and binding changes occurred in different trials within the same trial blocks. Error bars are standard errors.

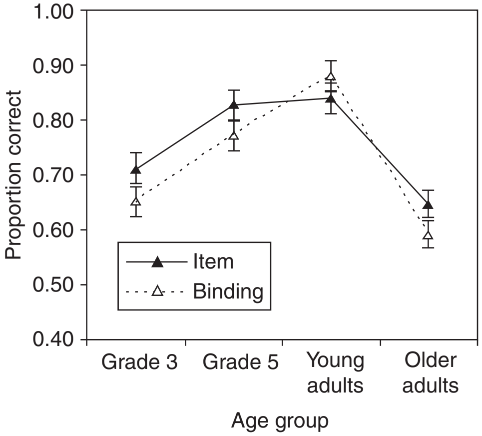

Fig. 10. From Cowan, Naveh-Benjamin, et al. (2006, Fig. 6). Proportion of trials correct for each age group (x-axis) in each type of trial block (graph parameter) within Experiment 2a. Item and binding changes occurred in different trial blocks, which allowed older adults to show no more binding deficit than children. Error bars are standard errors.

experiment in which item and binding changes were mixed together in the same trial block. One can see an inverted U-shaped developmental trend across the life span for item-change trials. Additionally, one can see that older adults did not do a good job of noticing the binding changes.

However, from those results it was not clear whether the older adults were incapable of noticing the binding changes or whether they simply did not attend to the correct information because the item changes were more salient. To address that question, Cowan, Naveh-Benjamin, et al. (2006, Experiment 2) developed a procedure in which item and binding changes appeared in separate trial blocks. The no-change control trials had to be adjusted accordingly. In the blocks with item changes, the cued item in no-change trials always was uniquely colored; in the blocks with binding changes, the cued item in no-change trials always had a color shared by another item in the array. Therefore, the participant could not guess whether the cued item had changed from the target array based on whether it was unique. The results are shown in Fig. 10. This study shows an inverted U-shape of developmental change in capacity, with little difference between children and older adults. A different pattern was found in an examination of bias; there was a monotonic trend to be less willing to indicate that a change had occurred. In terms of basic visual working-memory capacity, though, an inverted U-shape does appear to describe developmental change. Dividing attention with a tone identification task did produce lower capacity estimates in young adults, but did not reflect the other developmental trends that were unique to adult aging.

More work on aging has examined verbal working memory. Central to that work is the question of whether older adults' associative deficit will impede chunk formation, which should lead to smaller chunks than in young adults. It is critical to answer this question in order to address adequately the question of whether verbal working memory changes with adult aging. Allen and Coyne (1988) visually presented young adults and older adults with meaningless strings of letters. In this task, chunk size was externally manipulated by spacing between letters, with participants serially recalling between four and six letters when cued. For example, if six letters were to be recalled, the letter string was presented in the form "WRQB KX," influencing the formation of two chunks of different sizes. To examine how chunks were organized across age groups, error probabilities were computed for within-chunk and between-chunk boundaries; ideally, recall error should be lowest within a chunk, but higher between two chunks (Allen & Coyne, 1988). Although older adults recalled a fewer number of letter strings, there were no significant differences between age groups regarding the way that chunks were qualitatively organized (in terms of both number and size), suggesting that any deficits in memory were not due to organization in immediate memory. Similar studies confirmed these results, even when there was no spacing between letters to influence the number and size of chunks formed (Allen & Coyne, 1989; Allen & Crozier, 1992). Older adults are especially helped when stimuli in the environment support and influence recall, association, or grouping of items (Craik, 1983, 1986; Hay & Jacoby, 1999; Naveh-Benjamin, Craik, Guez, & Kreuger, 2005). Use of environmental support can greatly improve estimates in chunk size. In one study (Taub, 1974), older adults presented with letter strings that made words formed a greater number of chunks and had better recall than when the information was meaningless.

Naveh-Benjamin, Cowan, Kilb, and Chen (2007) examined chunking and capacity limits in working memory by presenting young adults and older adults with list of learned word pairs and lists of learned singletons (still presented in pairs within the list) to be recalled in serial order. This was the procedure of Cowan et al. (2004) in which, during a training phase, all words to be learned by participants received 4 different exposures, of which 0, 1, 2, or 4 exposures were as consistent pairs as opposed to singletons. There were also nonstudied control words. Several means were used to score the results so as to estimate chunks. In one simple method, a distinction was made between the number of pairs accessed and pair completion. Access referred to the number of pairs for which at least one item was recalled and pair completion referred to the proportion of accessed pairs for which both items were recalled. Figure 11 shows that there was a large, consistent difference in pair access favoring younger adults. Moreover, younger adults

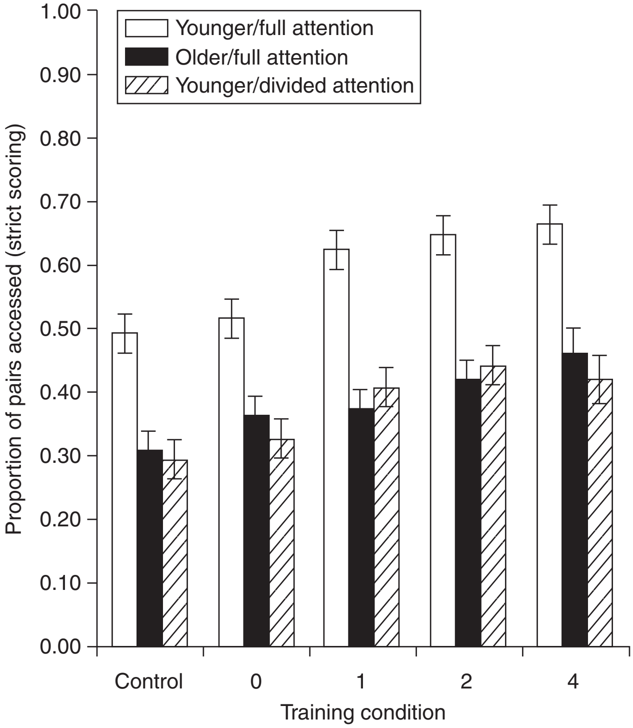

Fig. 11. From Naveh-Benjamin et al. (2007, Fig. 3). Proportion of presented pairs within a list to be recalled that were accessed in each pair-training condition. Access refers to at least one member of the pair being recalled. Older adults recalled fewer than young adults, but young adults under divided attention performed like older adults. Error bars are standard errors.

carrying out a divided attention task (tone identification) yielded pair access scores quite similar to older adults without divided attention. This result suggests that there is an adult aging decrement in capacity and that this decrement is attention related. A similar, clear result was obtained using various other estimates of chunk capacity. In contrast, there were only small and somewhat inconsistent group differences in the average sizes of chunks. It showed up in the 0- and 2-pairing conditions only. Across modalities and studies, then, adult aging sometimes results in a deficit in binding and association in working memory tasks but, regardless of those effects, consistently results in a deficit in some attention-dependent form of working memory capacity expressed in chunks.

There also has been considerable work emphasizing the reduction with adult aging in energy or attention control, as opposed to space or chunk capacity. Hasher and Zacks (1988) proposed that older adults have difficulty inhibiting irrelevant items, which allows them to take up valuable space in working memory. One type of evidence for this theory is comparable to the work on individual differences in proactive interference (e.g., Bunting, 2006; Conway & Engle, 1994). May, Hasher, and Kane (1999; also Lustig, May, & Hasher, 2001) used reading span and backward digit span tasks in two ways: with the usual, ascending order in which the list lengths get progressively longer across the session, and with a descending order in which the lists start at a relatively long length and get progressively shorter. The rationale is that ascending presentation allows considerable proactive interference to build up from one trial to the next by the time the difficult trials are presented. In the descending order, this was not the case. The deficit of older adults was eliminated with a descending presentation. One can, however, question the basis of this effect. Perhaps items in the first few trials are memorized and items only have to be held in a capacity-limited working memory store after a certain amount of proactive interference is present. Therefore, it is not clear whether the adult aging effect occurs because of space or energy differences. The same might be said of other results suggesting a resource limitation in child development or aging (e.g., Zelazo, Craik, & Booth, 2004). These could entail space limits, energy limits, or a combination of these limits.

A lot of work has documented that the speed of processing slows down in old age (Salthouse, 1996) and, as suggested above, this can have an effect on working memory performance. However, the literature still seems ambiguous about the fundamental causes of adult aging deficits in working memory. A speed deficit can increase the amount of energy needed to refresh representations in working memory, by making it necessary to remember more items at once rather than rotating among them rapidly. A faster decay rate requires a faster refreshing rate to rotate among items, making it more likely that items would be kept in an attention-based storage mechanism rather than successfully rotated. Energy limits and space limits may trade off, or insufficient space may have to be counteracted with an investment of energy. To some extent, these criticisms apply to all areas of the field; for example, Engle, Cantor, and Carullo (1992) found that an experimenter-paced version of the operation span task correlated with a cognitive measure (reading comprehension) whereas a subject-paced version did not. The capacity difference was critical only under speed pressure. Thus, the difficulty of separating causes of working memory limitations cannot be ignored.

Basak and Verhaeghen (2003) can be brought up as an example of how these difficult questions may be approached. They studied the range of subitizing, or apprehending small numbers of items without counting them. The classic literature shows that people can rapidly subitize about 4 items, after which a slower counting process is necessary (e.g., Mandler & Shebo, 1982). Recall that Tuholski et al. (2001) found no substantial relation between the range of subitizing and working memory in young adults. In contrast, Basak and Verhaeghen did find a difference between older and younger adults in the subitizing range; by their methods, younger adults were close to 3 items in that range, whereas older adults were close to 2 items. However, within the subitizing range, there was no difference in subitizing speed between younger and older adults. Rather than a general speed deficit (Salthouse, 1996), the results can be interpreted as showing that speed deficits in aging adults result from other deficits (in this case, engaging in slower counting processes for some lists that are more quickly subitized by younger adults).

E. Psychopathology

There is a vast literature on working memory deficits in diverse types of psychopathology, and it is beyond the scope of this chapter to review them. However, in this section we would like to point out that the recent literature has stressed the importance of energy (resource) limits in psychopathology, with emphasis on central executive processes and the frontal lobes (for reviews, see Engle, Sedek, von Hecker, & McIntosh, 2005). Here we issue a caution that this view may well be incomplete. An excellent case in point is provided by Gold et al. (2006). Their task was to compare two arrays of objects but some objects were irrelevant to the task (e.g., remember the orientations of red bars but ignore blue bars). Most of the time, the participants were tested on memory of one type of stimulus (e.g., the red bars) but, on occasional trials, they were tested on memory of the other type (e.g., the blue bars). In a separate test session, either type of bar could be relevant and each type was tested equally often. This test yielded evidence of how well attention was controlled: in particular, the steepness of the performance slope distinguishing between frequent, neutral, and infrequent conditions, with the highest performance for the frequent conditions and the infrequent feature attentionally filtered out. Better control results in attention more tightly focused on the frequent, most-relevant stimulus type and therefore better performance on it and a steeper slope when stimulus types are compared. The test also yielded evidence of how much visual working memory could hold: in particular, the sum of capacities observed in the frequent and infrequent conditions (e.g., red + blue bars in memory). Presumably, the number of items with the frequent feature plus the number with the infrequent feature make up the total contents of task-related working memory. As it happens, the total contents were less in schizophrenic patients than in normal control participants, but with little evidence of a change in attentional control (filtering). This finding is illustrated in Fig. 12.

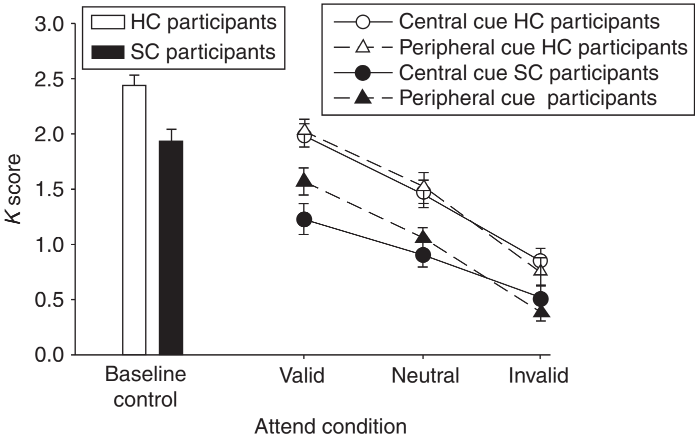

Fig. 12. From Gold et al. (2006, Fig. 2). HC = healthy control and SC = schizophrenic patient. The cue to look at colored squares on a particular side of the screen was valid on 60% of all trials, and a neutral cue was presented 20% of the time. The schizophrenic deficit showed up primarily in terms of capacity (lower intercept in patients) with only a small tendency for filtering out of invalid items to be affected (shallower slope in patients). Error bars are standard errors.

V. Addressing the Holistic-Versus-Analytic Distinction

We can surmise from the research on capacity that we have summarized that different levels of analysis lead to different measures for different purposes. If one is interested in what humans can accomplish, one finds that they can repeat up to seven or so verbal items (Miller, 1956); that the nature of the stimuli, such as how well it lends itself to verbal rehearsal, influences how well it can be recalled (Baddeley, 1986); and that people can learn to recall amazing amounts of information by grouping items together to form larger-level chunks (Miller, 1956) and data structures involving complex sets of associative relationships (Ericsson & Kintsch, 1995). However, that understanding is akin to knowing what an automobile does. If one wants to understand how an automobile works, or what makes one automobile more powerful or more efficient than another, one must adopt a more analytic framework in which one investigates how combustion engines work, how gear-shifting mechanisms work, and so on.

Research on capacity limits in chunks is of interest for at least two reasons. The first reason is that, as we have shown, there appears to be a simple limit of 3 to 4 chunks that, on average, occurs for adults across a large variety of test situations in which the contributions of rehearsal and other mnemonic strategies have been minimized. Knowing that limit can be of considerable use in predicting results in various new situations. The second reason is that, as we have further shown, the capacity limit appears to be a general limit that applies for items regardless of the modality in which they were presented, and for multimodal arrays of items; it is a limit vulnerable to distraction even from stimuli very different from the memoranda. This is of interest because it is what we would expect if what we are observing is the use of the focus of attention as a working-memory storage device. Given that this attention-related component of working memory is thought to be closely related to conscious awareness (e.g., Baars & Franklin, 2003; Cowan, 1995; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006), it may be that we are on the way to answering a fundamental philosophical question regarding how much information can be present in conscious awareness at once.

There are, however, multiple levels of analysis. After distinguishing between overall working memory ability and its division into components, an even finer level of analysis can be described. In specific, evidence of space (chunk capacity) limits can be further questioned as to their ultimate source. Attention-demanding, central executive processes presumably can be used to refresh items in memory (e.g., Cowan, 1992; Cowan et al., 1998; Hulme, Newton, Cowan, Stuart, & Brown, 1999; Raye, Johnson, Mitchell, Greene, & Johnson, 2007) and, at the macroscopic level, the process of doing so could look the same as using attention to hold the items concurrently. Therefore, at a microscopic level, we are still not sure what is going on and it is only at the intermediate level of analysis that we can draw conclusions about working memory capacity in chunks. At that intermediate level of analysis, the evidence is strong that (1) there is such a basic limit, (2) it is important in information processing and cognition, and (3) it appears to be important in understanding ideographic differences, including individual differences among young adults, life span developmental differences, and differences due to psychopathology.

VI. Conclusion

We have considered the concept of working memory according to separate potential contributions of space (i.e., chunk capacity), time (i.e., decay and speed factors), and energy (i.e., resources). The emphasis was in showing that the notion of chunk capacity is not intractable, can be separated from other factors, and plays an important role in both nomothetic and ideographic considerations of working memory according to an analytic perspective. Miller (1989) explained how he (Miller, 1956) was using only the magical number seven as a rhetorical device, and little more than a joke, to organize his presentation of several otherwise unrelated research areas. Although the approach in that classic paper may have set a skeptical trend in the field, we, after Broadbent (1975), believe that there are some important generalizations to be had regarding fixed capacity limits. The difficulty in reaching those generalizations is in identifying the chunks that individuals use and we have presented methods to examine what the chunks are. What appears to be a general capacity limit in working memory (Cowan & Morey, 2007; Saults & Cowan, 2007) is closely related to the contents of the conscious mind. It has the advantage of being easier to quantify than the resource limits that also appear to exist and to differ among individuals (e.g., Kane et al., 2004). The space and energy limits may occur for a common reason if, in fact, they share a focus of attention that can zoom out to apprehend a field of objects or zoom in to hold on to a goal in the face of a prepotent response going against that goal. With the devotion of enough time and resources, we will learn the truth about chunk capacity and its space metaphor.

REFERENCES

Allen, P. A., & Coyne, A. C. (1988). Age differences in primary organization or processing variability? Part I: An examination of age and primary organization. *Experimental Aging Research, 14*, 143–149.

Allen, P. A., & Coyne, A. C. (1989). Are there age differences in chunking. *Journal of Gerontology, Psychological Sciences, 44*, P181–P183.

Allen, P. A., & Crozier, L. C. (1992). Age and ideal chunk size. *Journal of Gerontology Psychological Sciences, 47*, P47–P51.

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General, 135*, 298–313.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111.

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology, 45*, 153–219.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*, 622–628.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5*, 119–126.

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

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

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in cognitive sciences, 4*, 417–423.

Baddeley, A. (2001). The magic number and the episodic buffer. *Behavioral and Brain Sciences, 24*, 117–118.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

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

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

Baddeley, A. 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.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*, 176–189.

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.

Basak, C., & Verhaeghen, P. (2003). Subitizing speed, subitizing range, counting speed, the Stroop effect, and aging: Capacity differences and speed equivalence. *Psychology & Aging, 18*, 240–249.

Binet, A., & Simon, T. (1916/1980). *The development of intelligence in children*. Nashville, TN: Williams Printing Company.

Bleckley, M. K., Durso, F. T., Crutchfield, J. M., Engle, R. W., & Khanna, M. (2003). Individual differences in working memory capacity predict visual attention allocation. *Psychonomic Bulletin & Review, 10*, 884–889.

Bolton, T. L. (1892). The growth of memory in school children. *American Journal of Psychology, 4*, 362–380.

Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.

Broadbent, D. E. (1975). The magic number seven after fifteen years. In A. Kennedy and A. Wilkes (Eds.), *Studies in long-term memory* (pp. 3–18). Oxford, England: John Wiley & Sons.

Bunting, M. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 183–196.

Burtis, P. J. (1982). Capacity increase and chunking in the development of short-term memory. *Journal of Experimental Child Psychology, 34*, 387–413.

Case, R. (1972). Validation of a neo-Piagetian mental capacity construct. *Journal of Experimental Child Psychology, 14*, 287–302.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*, 403–416.

Chen, Z., & Cowan, N. (2005). Chunk limits and length limits in immediate recall: A reconciliation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1235–1249.

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.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8*, 331–335.

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

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

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

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford Psychology Series #26. New York: Oxford University Press.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press.

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*. Hove, East Sussex, UK: Psychology Press.

Cowan, N., & AuBuchon, A. M. (2008). Short-term memory loss over time without retroactive stimulus interference. *Psychonomic Bulletin & Review, 15*, 230–235.

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., et al. (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. (2006). Rethinking speed theories of cognitive development: Increasing the rate of recall without affecting accuracy. *Psychological Science, 17*, 67–73.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34*, 1754–1768.

Cowan, N., & Morey, C. C. (2007). How can dual-task working memory retention limits be investigated? *Psychological Science, 18*, 686–688.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult. *Developmental Psychology, 42*, 1089–1102.

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., Saults, J. S., & Morey, C. C. (2006). Development of working memory for verbal-spatial associations. *Journal of Memory and Language, 55*, 274–289.

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.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London, 302*, 341–359.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix and H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 409–421). New York: Elsevier Science Publishers.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarman, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112*, 3–42.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*, 205–211.

Delaney, P. F., & Sahakyan, L. (2007). Unexpected costs of high working memory capacity following directed forgetting and contextual change manipulations. *Memory & Cognition, 35*, 1074–1083.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 972–992.

Engle, R. W., Sedek, G., von Hecker, U., & McIntosh, D. N. (Eds.) (2005). *Cognitive limitations in aging and psychopathology*. New York: Cambridge University Press.

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: General, 128*, 309–331.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208*, 1181–1182.

Ericsson, K. A., Delaney, P. F., Weaver, G., & Mahadevan, R. (2004). Uncovering the structure of a memorist's superior "basic" memory capacity. *Cognitive Psychology, 49*, 191–237.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

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

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning & Verbal Behavior, 5*, 351–360.

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., Lane, P. C. R., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*, 236–243.

Gold, J. M., Fuller, R. L., Robinson, B. M., McMahon, R. P., Braun, E. L., & Luck, S. J. (2006). Intact attentional control of working memory encoding in schizophrenia. *Journal of Abnormal Psychology, 115*, 658–673.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*, 316–322.

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen and F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5, pp. 500–639). New York: Academic Press.

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., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process. *Psychological Science, 16*, 70–76.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences, 11*, 236–242.

Hambrick, D. Z., & Engle, R. W. (2001). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*, 339–387.

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

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology & Aging, 14*, 122–134.

Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General, 136*, 217–240.

Hulme, C., & Muir, C. (1985). Developmental changes in speech rate and memory span: A causal relationship. *British Journal of Developmental Psychology, 3*, 175–181.

Hulme, C., Newton, P., Cowan, N., Stuart, G., & Brown, G. (1999). Think before you speak: Pause, memory search and trace redintegration processes in verbal memory span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 447–463.

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

Jacobs, J. (1887). Experiments on "prehension." *Mind, 12*, 75–79.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*, 683–702.

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., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 336–358.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47–70.

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.

Kim, S.-Y., Kim, M.-S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Science (PNAS), 102*, 16524–16529.

Kynette, D., Kemper, S., Norman, S., & Cheung, H. (1990). Adults' word recall and word repetition. *Experimental Aging Research, 16*, 117–121.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*, 339–354.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11*, 771–790.

Logie, R. H., & Gilhooly, K. J. (1998). *Working memory and thinking*. Hove, England: Psychology Press.

Lovett, M. C., Reder, L. M., & Lebière, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135–182). Cambridge: Cambridge University Press.

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.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111*, 1–22.

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

Massaro, D. W. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology, 83*, 32–39.

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.

Miller, G. A. (1989). George A. Miller. In L. Gardner (Ed.), *A history of psychology in autobiography* (Vol. VIII, pp. 391–418). Stanford, CA: Stanford University Press.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston, Inc.

Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., & D'Esposito, M. (2000). Aging and reflective processes of working memory: Binding and test load deficits. *Psychology and Aging, 15*, 527–541.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

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.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*, 1353–1380.

Mulligan, N. W., & Lozito, J. P. (2007). Order information and free recall: Evaluating the item-order hypothesis. *Quarterly Journal of Experimental Psychology, 60*, 732–751.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1170–1187.

Naveh-Benjamin, M., Cowan, N., Kilb, A., & Chen, Z. (2007). Age-related differences in immediate serial recall: Dissociating chunk formation and capacity. *Memory & Cognition, 35*, 724–737.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*, 520–537.

Naveh-Benjamin, M., Hussain, Z., Guez, J., & Bar-On, M. (2003). Adult age differences in episodic memory: Further support for an associative-deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 826–837.

Neath, I., Bireta, T. J., & Surprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review, 10*, 430–434.

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.

Pascual-Leone, J. (2005). A neoPiagetian view of developmental intelligence. In O. Wilhelm and R. W. Engle (Eds.), *Understanding and measuring intelligence* (pp. 177–201). London: Sage.

Pascual-Leone, J. A. (1970). Mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*, 301–345.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*, 369–378.

Posner, M. I., & Konick, A. F. (1966). On the role of interference in short-term retention. *Journal of Experimental Psychology, 72*, 221–231.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex, 43*, 135–145.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*, 211–227.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39*, 418–436.

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74*, 1807–1822.

Saito, S., & Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in reading span performance. *Journal of Memory and Language, 50*, 425–443.

Salthouse, T. A. (1985). *A theory of cognitive aging*. Amsterdam: North-Holland.

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

Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General, 136*, 663–684.

Saxe, M. D., Malleret, G., Vronskaya, S., Mendez, I., Garcia, A. D., Sofroniew, M. V., et al. (2007). Paradoxical influence of hippocampal neurogenesis on working memory. *Proceedings of the National Academy of Sciences (PNAS), 104*, 4642–4646.

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

Shah, P., & Miyake, A. (2005). *The Cambridge handbook of visuospatial thinking*. New York: Cambridge University Press.

Slak, S. (1970). Phonemic recoding of digital information. *Journal of Experimental Psychology, 86*, 398–406.

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

Stevanovski, B., & Jolicoeur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition, 15*, 532–563.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Taub, H. A. (1974). Coding for short-term memory as a function of age. *Journal of Genetic Psychology, 125*, 309–314.

Towse, J. N., Hitch, G. J., Hamilton, Z., Peacock, K., & Hutton, U. M. Z. (2005). Working memory period: The endurance of mental representations. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 58A*, 547–571.

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition, 34*, 1704–1719.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory & Cognition, 29*, 484–492.

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.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104–132.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*, 500–503.

Wechsler, D. (1944). *The measurement of adult intelligence* (3rd edn.). Baltimore, MD: The Williams and Wilkins Company.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children—third edition (WISC-III)*. San Antonio, TX: The Psychological Corporation.

Weiss, V. (1995). Memory span as the quantum of action of thought. *Cahiers de Psychologie Cognitive, 14*, 387–408.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131*, 48–64.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440*, 91–95.

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica, 115*, 167–183.