The Legend of the Magical Number Seven

Nelson Cowan, Candice C. Morey, and Zhijian Chen

University of Missouri

Address Correspondence to:

Nelson Cowan

Department of Psychological Sciences

University of Missouri

18 McAlester Hall

Columbia, MO 65211

USA

E-mail: CowanN@missouri.edu

Telephone: 573-882-4232

This is a draft of a chapter for an edited volume. The reference is:

Cowan, N., Morey, C.C., & Chen, Z. (in press). The legend of the magical number seven. In S. Della Sala (Ed.), *Tall tales about the brain: Things we think we know about the mind, but ain't so*. Oxford University Press.

.

Origin of the Legend of Seven

Individuals who know very little about experimental psychology are still likely to have heard or read that people can keep in mind about seven items. Telephone numbers were developed with some concern for people's ability to remember the numbers, and local calls in the United States typically require dialing seven digits (or, in some countries, just six digits). Intelligence test batteries include a test called *digit span* in which one is to repeat a list of random digits in the presented order; the digits in the list change from one trial to the next, and the length of the list keeps increasing every few trials until the tested individual cannot repeat any lists correctly. Normal adults typically can repeat lists of about seven digits. This maxim of seven has often been applied to daily life. For example, some self-help sources proclaim that a good oral presentation should include up to seven points on the outline. The number seven appears in dinner-party talk, along with other psychological folk wisdom such as the best way to raise children or how to bargain with salespeople effectively.

How did this information get established in the public mind? It goes back to a seminal journal article by George Miller[1] that was published in 1956, in the formative days of a new field that came to be known as *cognitive psychology*, the experimental study of thought processes such as memory, attention, imagery, and language comprehension and production. Miller's article was written in a very engaging and entertaining fashion, in part because it began as an hour-long conference presentation before it was molded into a written article. It begins with the author's humorous confession that he has been persecuted by the integer seven. He goes on to discuss three types of psychological task in which this number has emerged.

The first and most obvious task is *immediate memory*, such as the digit-span task or similar tasks in which lists are presented and must then be repeated without delay in the presented order. No matter whether the stimuli are words, letters, or digits, lists of only about seven of them can be recalled. This differs somewhat from one individual to the next and from one type of memoranda to the next and, indeed, the title of Miller's article included the phrase, "the magical number seven, plus or minus two."

In a second type of task that Miller discussed, *absolute judgment*, a single stimulus is presented and its correct label has to be recalled. This is tough when the stimuli are simple and differ in only one dimension, such as a series of lines of different lengths or a series of tones of different pitches, each with a different label. It turns out not to matter whether the stimuli differ only slightly or whether they differ a lot. So long as they differ enough that the research participants can see or hear the differences between them when they are placed side by side (or, for sounds, in close succession), the same memory limit applies. The task of identifying an isolated stimulus can be accomplished adequately only when there are no more than about seven stimulus choices, again varying depending on the exact context.

A third type of task that Miller discussed is the *span of attention*. In the relevant task, a set of haphazardly-arranged objects (or perhaps dots on a computer screen) must be *enumerated* as quickly as possible; that is, the participant must indicate how many objects are present. Now, your own experience probably tells you that enumerating the objects in a set of, say, two is a very different experience from enumerating objects in a larger set of, say, eleven. The two objects can be enumerated very quickly, on the basis of rapid recognition or attention to both at once, without counting. It is a different matter with eleven objects. One must carefully keep track of which ones have been counted while one is in the process of counting the others. Miller said that sets of up to about six or seven objects are enumerated rapidly whereas, with higher numbers, the time to give an answer begins to rise steeply with each added object in the set.

These three phenomena not only comprised an impressive display of evidence; they comprised evidence central to the newly developing field of cognitive psychology. In an earlier era, philosophically-oriented psychologists such as William James had pointed out that there were several types of memory. James[2] distinguished between the small amount of information that is or recently was in one's conscious mind, which he termed *primary memory*, and the large storehouse of knowledge that one collects over a lifetime, which he termed *secondary memory*. If cognitive psychology was to become scientific, though, there had to be a way to measure and characterize these types of memory. The estimate that about seven items could be held in primary memory would be a giant step toward that end. In the era when Miller wrote, psychologists from the behaviorist tradition, counter to James, were advising that one should study stimuli and responses only, and should avoid making statements about unobservable entities inside the human head such as memory or mental imagery. On the basis of Miller's article and other, converging work published around the same time, that sentiment was overturned for cognitive psychologists. Regarding Miller's findings, if people could recall about seven items, there must be some holding mechanism in the brain, corresponding to James' primary memory, that could hold about seven items at once but not much more. The well-described findings were repeated often by psychologists and they eventually reached the general public, in much the same way that concepts from Sigmund Freud earlier had reached the public.

The Intent Behind the Legend

There are aspects of Miller's 1956 article[1] that have left the careful reader with a bit of confusion regarding what he intended to say. He does not actually make the claim that memory span, absolute identification, and enumeration tasks call upon the same faculty of the mind limited to seven or so items. Instead, he ends with a note on the mystery of the convergence of many phenomena:

> “What about the seven-point rating scale, the seven categories for absolute judgment, the seven objects in the span of attention, and the seven digits in the span of immediate memory?...Perhaps there is something deep and profound behind all of these sevens, something just calling out for us to discover it. But I suspect that it is only a pernicious, Pythagorean coincidence.” (p. 96)

Often when one ponders a legend and learns more, the supporting evidence can be seen to have different implications than one might have thought according to the legend that developed. In this case, it turns out that Miller was not very interested scientifically in the number seven. Perhaps if he had been, he would not have attached the adjective "magical" to it. As he explained in an autobiographical essay[3], he was asked to give an hour-long presentation at a point in his career when he did not feel that he had any one research topic developed enough to take up that time period. He did, however, have some research on immediate memory and on absolute judgment. He did not want to give two unconnected reports of these research topics and at first saw no common theme between them. However, he then discovered that they shared the number seven in terms of research participants' limits in performance. He decided to make that limitation a theme of the talk to tie them together and, to add an air of legitimacy, threw in the research on enumeration. However, the reference to "plus or minus" seven was supposed to convey the humorous notion that a magical number could have a margin of error. This is an amazing way for a scientific legend to be born.

One concept that was more important to Miller[1] was the concept of *chunking*. This means taking multiple items and putting them together to form new groups or chunks. Before Miller, psychologists tried to measure information in *bits*, a term frequently used in computing, meaning a choice between two options. Two bits equals 2^2 or 4 options, three bits equals 2^3 or 8 options, and so on. For example, how many yes/no questions would it take you to guess which English letter a friend is thinking of? With your first question (eliciting one bit of information), you could ask if the letter comes before N in the alphabet, narrowing the choices down to 13 of 26 letters, or half of the alphabet; with your second question (eliciting a second bit), you could narrow the choices down to approximately half of that half; and so on, until you could determine which letter it was. There also is a mathematical definition of bits on a scale that includes fractions; without going into this definition, it is enough here to give the example that 2.6 bits is something more than 2 bits but smaller than 3 bits. If one considers digits from the set 0-9, there are ten choices so each digit conveys somewhere between 3 and 4 bits of information. If one considers the 26 English letters, each letter conveys somewhere between 4 bits (2^4 = 16 choices) and 5 bits (2^5 = 32 choices). There are many thousands of English words, so the bit measure for an English word would be considerably higher.

However, it turned out that bits did not matter for actual research participants. Memory span is about the same number of items when the items are random digits, random letters, or random words. It appears that immediate memory should be measured not in bits, but in units that are psychologically meaningful. Each meaningful unit is called a chunk. In this regard, human memory appears to operate in a manner quite different from computer memory, which is composed of many locations that can be turned on or off, each worth one bit of information.

One might have thought that bits would be important for humans, given that each nerve cell is in a firing or non-firing state at any moment and therefore may convey only 1 bit of information. Apparently, though, this binary property of individual nerve cells is not what is important for immediate-memory limits. Perhaps that is because large portions of the brain's memory system can participate in immediate memory; not just a relatively small, dedicated portion of the memory locations as in a computer. What may be important is limitations in the firing patterns that nerve cells can take on at any moment, such that only a few ideas can be actively represented concurrently.

Miller and one of his colleagues found that stimuli can be transformed in a way that makes them easier to remember, by reducing the number of chunks. In the binary numerical system that is used to encode computer memory locations using only the digits 0 and 1, the rightmost digit reflects how many ones there are, the next digit to the left reflects how many twos, and the next digit to the left of that reflects how many fours; so 001 = 1; 010 = 2; 011 = 3; 100 = 4; 101 = 5; 110 = 6; and 111 = 7. It would be difficult to remember the binary string 011-111-101-110, yet much easier to remember the familiar decimal numerical equivalent, 3-7-5-6. If one knows the binary system, one can recode the binary string into its decimal equivalent. In the example given here, recoding reduces the load on immediate memory from 12 chunks (the binary digits shown) down to only 4 chunks (the digits 3, 7, 5, and 6). Another example that makes the concept clear is memorization of the letter string *USAFBICIA*. This looks like 9 chunks (single, unrelated letters) but they can be reduced to three acronyms: USA (United States of America), FBI (Federal Bureau of Investigation), and CIA (Central Intelligence Agency). For someone who knows these acronyms by heart and notices these patterns, there are only 3 chunks to be remembered.

In sum, it was not the number seven per se that fascinated Miller, but rather the processes that were used to encode information and the nature of the units that were meaningfully encoded. This was intimated in the tone of the closing comments in his 1956 article[1] and was made clear in his later autobiographical discussion[3]. People could recall about seven chunks, regardless of the processes that were involved in deriving those chunks from the stimuli to be recalled.

The formation of chunks in immediate or primary memory often made use not only of the information present to the research participant, but also of prior knowledge that was already present in long-term or secondary memory. It is worth noting that there have been demonstrations that practically anything can be held in immediate memory, if there is enough knowledge to back it up. Anders Ericsson[4] and colleagues trained an individual to increase his digit span from the usual seven or so up to 80 digits, in the course of a year. This individual was an athlete who already had memorized many record running times. This made it easier to transform digits into multi-digit chunks. For example, 3.98 might be the record time in minutes to run a mile on a certain type of track. This could be supplemented with new chunks, such as 85.7 as the age of a pretty old man. Grouping sets of three and four digits together to form new chunks, over a period of months this special individual (or was he just specially motivated?) learned to repeat lists of about 20 digits, presumably organized into 5 to 7 larger chunks. Then, somehow he learned to combine several chunks into even larger super-chunks, so that he eventually could repeat series of about 80 digits. This skill did not generalize; his memory for letters or words remained at about seven.

Similarly, Jeffrey Rouder and colleagues[5] recently found that, with extended practice, absolute judgments for line lengths could be extended considerably beyond the seven or so distinct labels that Miller noted. We do not know just how chunking is involved in absolute judgments but one possibility is that there is a limit in how many categories can be kept distinctly in mind during the test, which might be overcome through extended familiarity with the categories.

Problems With The Number Seven

There were findings resulting in seven or so items remembered, and these findings require some explanation. Still, one might question whether seven actually is a fundamental number of immediate memory. Consider this. If people are able to perceive multiple items in terms of chunks that they already know (such as the acronym IRS) might it not also be possible for them to form new chunks rapidly? Why is it, for example, that the seven digits in a telephone number are typically presented in two groups, in the form # # # - # # # #? It seems reasonable to suppose that some rapid grouping process goes on to ease the process of recall by reducing the number of independent units that have to be recalled. These questions did not get a great deal of immediate attention, however. One reason was that, after 1956, George Miller's career seemed to veer more into the study of language and categorization, as opposed to primary memory.

Published just four years after Miller's famous article, a 1960 article by George Sperling[6] became another lasting classic in the field of cognitive psychology and yielded a different answer about primary memory. The study's main point was that a large amount of information about how a visual stimulus looks is stored in the mind for a very short time, but the study also provided information about primary memory. On each trial, a spatial array of characters (such as letters) was flashed on the screen briefly. The task was to record all of the characters in the array, or some part of the array. A large amount of elegant experimental work was included in the article. It was found that if the row of the array to write down was indicated by a tone presented quickly enough, before sensory memory had faded, it was possible to write down most of the characters in that row. This showed that sensory memory could hold visual information from at least 12 characters at once. However, if there was no tone cue and the entire array had to be written down, there was a more severe limit in performance so that only about 4 of the items could be written down. The theoretical model for this task was that information had to be processed, from a visual form in sensory memory into a more categorized or labeled form in primary memory, before it could be reported. Either primary memory could hold only about 4 items, or sensory memory did not last long enough to allow more than 4 items to be processed. One can imagine an analogy in which a painter must paint objects onto a canvas of limited size (like primary memory) using an open tray of paint that is plentiful but dries up extremely rapidly (like a fading sensory memory). The number of objects that can be painted onto the canvas depends on both the size of the canvas and the time available before the paint becomes too dry to use. We will return to this issue later.

There also were studies indicating that people could recall roughly 4 clusters or chunks of objects, though experts could recall chunks comprising more objects. This research involved people's ability to recall the pieces on a chessboard, as a function of their expertise in chess.[7] Work continuing along this line[8] has suggested that even the notion of a chunk is often an oversimplification for what can be a broad network of associations between items, or template.

There were a few studies by other investigators looking at the issue of grouping in immediate recall. For example, Tulving and Patkau[9] carried out a study in which people were asked to remember strings of 24 words that were in jumbled order, or that resembled coherent English to varying degrees (for example, "The best grain stamps made in America you beast that see something..."), or that were perfectly coherent English sentences. The task was to recall the words in any order (*free recall*). Whenever runs of several words were recalled in the same order in which they were presented, each such run counted as a single chunk. Many more words were recalled in the sequences that were better approximations to English, but the measured number of chunks recalled remained fixed across conditions, at 4 to 6 chunks. It was just that more coherent strings of words led to *larger* chunks recalled, not *more* chunks. Other methods were invented in attempts to identify chunks clearly, such as making the assumption that the task of recalling lists in order (*serial recall*) would proceed relatively smoothly within a chunk but would be more likely to encounter difficulty between chunks.[10] Overall, though, the magical number seven was neither seriously questioned nor put to many stringent tests in the early days. Some investigators lived by it, and others probably were skeptical and ignored it, perhaps taking their cues from the ending of Miller's article in which it was said that the magical number seven was probably just a coincidence.

The year 1975 was, in hindsight, an important one for the study of immediate memory. By this year, the magical number seven had been recognized as a classic finding that had withstood the test of time. Yet, two papers were published that also have had a lasting impact and have cast doubt on the magic of the number seven.

First, Alan Baddeley and colleagues[11] showed that it is not simply the number of meaningful units that mattered in immediate recall; word length mattered. Lists of words that took longer to pronounce were not recalled as well as lists of the same number of words that could be pronounced more quickly. The explanation of that finding was that people refresh their verbal memories by rehearsing the words (that is, imagining saying the words to themselves), a process that can be carried out more efficiently for short words. If the entire list were rehearsed over and over, for example, the time between one rehearsal of a particular word and the next rehearsal of the same word would be shorter if the words were shorter, leaving less time for forgetting. It may be that rehearsal takes place in a more complex or piecemeal manner than that but, in any case, many such methods of rehearsal would lead to the expectation of the word-length effect that actually was obtained. Baddeley has amassed a large amount of information about primary memory in subsequent work, and a time-related limit remains an important part of the theorization that has become predominant in the field of what is now called *working memory*, or primary memory as it is used to help do work such as solving problems and comprehending and producing language.

Second, rehearsal aside, in a 1975 book chapter[12] one of the founding fathers of the field of cognitive psychology, Donald Broadbent, began to question how fundamental the number seven actually was in primary memory. The logic of this challenge was similar to what has been stated above. It was pointed out that although people typically could remember up to about seven items, perhaps a more meaningful number was the number of items that people could remember flawlessly (because presumably those items are recalled without relying on a mental strategy that can fail). For sets of only three items, memory was nearly flawless. Adding a fourth or fifth item resulted in a set that could usually be recalled correctly; adding more made the situation worse. It therefore appeared that three was a basic capacity limit and that rehearsal, grouping, or other strategies or mental tricks might be used sometimes to increase the number recalled beyond that basic capacity. As analogies for these strategies, a juggler can keep multiple balls off the ground by repeatedly renewing their upward momentum (like rehearsing), and a person can keep multiple balls off the ground by putting several of them together on a plate (like chunking). However, jugglers sometimes make mistakes and balls sometimes roll off of plates. Broadbent pointed out other phenomena to support the notion that the magical number was not seven, but three. For example, when one attempts to recall items from a category in secondary memory, one tends to recall in bursts of three items. Try, for example, to name countries of the world as quickly as possible and you will notice that they tend to be produced in spurts of just several countries at a time.

Is There a Magical Number After All?

Much more recently, one of the present authors (Nelson Cowan) wrote a literature review[13] that examined Broadbent's hypothesis more broadly and systematically. It suggested that, across many types of experiment, something like a semi-magical number 4 (plus or minus two, varying across individuals and situations) actually exists. To find this result, one must include only procedures in which the items are well known and in which it is impossible to form larger chunks from the items. This can be accomplished, for example, by presenting many items in an array, like Sperling[6], with the array presented only briefly so that there is not enough time to think about all of the items in a way leading to extensive chunking. In a particularly compelling demonstration of this, called multi-object tracking[14], there are multiple objects on the computer screen and then several of them momentarily are marked to stand out (for example by flashing). When this stops, so that all the items look alike again, they wander around the screen randomly, in different directions. When they stop, the research participant is quizzed regarding whether a certain object was one of the previously-marked objects, or not. People typically can follow or track a maximum of 4 objects, and sometimes fewer.

Formation of new chunks also can be prevented by presenting lists of spoken items in a situation in which attention is diverted to another task at the time that these items are presented, making rehearsal impossible. Then the spoken items have to be recovered from the stream of auditory sensory memory when a cue to recall them is presented, just after the list in question has ended. If chunking is not possible, it is assumed that each item remains a single chunk in primary memory. Under such circumstances, about 4 items (that is, presumably, single-item chunks) can be recalled. Similar results are obtained if the spoken items are attended but covert verbal rehearsal is prevented by requiring that the participant at the same time repeats a meaningless phrase during the testing, a procedure known as *articulatory suppression*.

Could it be shown that this capacity limit of about 4 chunks, observed in so many circumstances when chunks were presumably limited to one item each[13], applies also when chunking is possible? If so, then this capacity limit will gain considerable generality. This does seem to be the case with some of the previous results[7, 9]. However, the question has so rarely been studied that it cannot be considered to have been decided.

The reason for the limit of about 4 chunks also has not been determined. One reason it could occur is that the chunks have to be held in the focus of attention, which is limited in capacity. Another possibility is that the chunks do not have to be held in a region of the mind that is limited in capacity, but that the chunks interfere with each other if they include similar features or concepts.

In the final section of this chapter, we will illustrate the ongoing controversy and how it might be resolved in the future, by reporting on some recent work on capacity limits.

Some Recent Studies on Immediate-Memory Capacity Limits

Recently, work has been conducted to help ascertain that there really are capacity limits in immediate memory, and also to determine what the reasons for capacity limits might be.

One recent study conducted to ascertain the capacity limit[15] went back to the standard technique to study immediate memory that was discussed by Miller[1], serial recall. Instead of preventing chunking (as in Cowan's[13] previous approach), steps were taken to control chunking. In a training session that preceded serial recall, words were presented either singly or in pairs. Each word was presented 4 times but a proportion of those presentations involved consistent pairs of words. For example, within the training sequence of words, the words *brick* and *hat* might each be presented twice by themselves, and twice in the consistent pair *brick-hat*. This mixture was termed the 2-pairing condition. The different training conditions (the 0-, 1-, 2-, and 4-pairing conditions) used with different words are outlined in Figure 1. The 0-pairing condition involved no training with word pairs per se whereas, at the other end of the continuum of training conditions, the 4-pairing condition involved consistent training with words in pairs. The expectation was that more frequent pairing would increase the likelihood that the pair would be remembered as a single chunk in serial recall, rather than as two separate words. There also was a cued-recall test in which, for example, the word *brick* was presented and the correct response was *hat*, if that was a pair that had been presented.

- - - - - Figure 1 here - - - - -

To encourage the recall of learned pairs, items within the 8-word lists to be recalled were presented in pairs. Each list included words from a single training condition. Each list that was composed from words in the 1-, 2-, or 4-pairing condition included only pairs that were already familiar from training. For example, somewhere within an 8-word list of words from the 2-pairing condition, the pair *brick-hat* would appear, if it happened to be part of this training condition for a certain participant (as in the example above). Results of this experiment are depicted in Figure 2. The blue triangles show that the number of words recalled in the correct list positions increased markedly as a result of more paired training.

The red circles in Figure 2 show the number of chunks recalled, using one of several measures of chunking. This reflects the sum of 1-word chunks, or singletons, and 2-word chunks, or learned pairs. (Several methods were used to ascertain which pairs had been learned.) The clear finding was that the number of chunks recalled stayed constant across learning conditions, at an average of about three and a half chunks, even though the number of words recalled increased with pair training.

- - - - - Figure 2 here - - - - -

Another, very different research procedure[16] will now be introduced, not only to show the variety of procedures leading to a capacity limit, but also to permit a discussion of some recent research on the question of why the capacity limit occurs. In this procedure, a haphazard array of colored squares is briefly presented and is followed, after a short break of up to a half second, by another array of squares that is identical to the first one or differs in just the color of one of the squares. The task is to indicate whether the array has changed or not; half the time, the correct answer is "yes" and half the time it is "no." To make the decision easier, a circle can appear surrounding one square, the participant having been instructed that, if anything changed, it was the color of the circled square. The procedure is illustrated in Figure 3.

- - - - - Figure 3 here - - - - -

This task is easy with up to 4 squares in the array, but it becomes progressively harder as the number of squares in the array (called the *set size*) increases beyond 4. There is a way to use the results of the experiment to estimate the number of squares from the first array that had to be held in primary memory, taking into account guessing (Cowan[13], p. 166). For all set sizes, the estimate comes out to be about three and a half items. If we can assume that the arrays are flashed too briefly for multi-item chunks to be formed, this means an average of three and a half chunks.

This array-comparison procedure may be helpful in understanding capacity limits and what factors cause them, because it is a nonverbal procedure. In a verbal procedure, the process of rehearsal may get in the way of understanding the fundamental capacity limit, as discussed above. In a nonverbal procedure, as we will show, this can be less of an issue.

Recall that one explanation for the capacity limit is that some information in primary memory must be held in the focus of attention, as William James[2] implied in his writing long ago. It is clear that the focus of attention is limited; perhaps it is the focus of attention that has a capacity of three or four chunks of information. To examine this possibility, one recent study[17] used a dual task in which a spoken list of digits was to be retained and recited aloud during the reception and retention of the first array on the trial. Given that spoken digits and visual arrays have very different features, they need not interfere with one another unless both of them require the same resource that is severely limited in capacity, such as the focus of attention and its potential ability to hold information.

There were four different conditions. In one condition, there was no digit recitation. In two memory load conditions, a random two- or a seven-digit number had to be recited. The fourth condition was a control to make sure that it was not recitation per se that hurt recall. In this condition, it was the participant's own seven-digit telephone number that had to be recited during the trial. Because the number was known, it did not impose a load on primary memory. However, it involved digit recitation comparable to the seven-digit load condition. Thus, it was only the seven-random-digit condition that imposed the kind of load that should make demands on the focus of attention, in addition to articulation.

The results of this study are shown in Figure 4 in terms of the estimated capacity in each condition (averaged across different array sizes). As expected according to the theory that the capacity limit is in the focus of attention, performance was impaired by the seven-digit memory load, but not by the other recitation conditions. The effect of the memory load was especially great when the load was recited incorrectly, in which case the valiant attempts to retrieve the verbal information were probably distracting in and of themselves.

- - - - - Figure 4 here - - - - -

Recent studies differ in their conclusions. In one study[18], the interference between visual arrays and digit lists was considerably less than in the study shown in Figure 4. One possibly important difference between the studies was that only the study showing more interference[17] required that the digit memory load be recited aloud during the presentation of visual arrays. Other recent work suggests that the retention of verbal information can require attention even if it is to be held silently, provided that two conditions are met. The information must be beyond the amount that can be conveniently and silently rehearsed, yet it must be unstructured enough that it cannot be greatly simplified or chunked using information from long-term memory[19].

Other recent research has tied the visual array procedure to neural functioning. Individuals with a larger capacity for the colored squares appear to show electrical signals emanating from the brain that increase more as the number of squares per array increases from two to four[20]. Images of neural responses to stimulation based on functional magnetic resonance imaging (fMRI) show select areas of the brain that respond in a manner similar to the capacity limits observed in behavioral work[21].

There are many different experimental procedures and each one has to be analyzed carefully before we will know whether a similar "magical number" truly applies to all of them, and for the same reason. In one sort of procedure, a visually-presented list of words is followed by a probe word, which has to be judged to be present in the list or absent from it. The reaction time to the last word in the list is shorter than the reaction time to the other words, leading to the possible conclusion that, actually, only one item is held in the focus of attention in such situations[22]. If this is the case, then the capacity limit of three to four items might not apply to such situations. However, further work has shown that the fast reaction time spreads from one item to four items as the participants become highly practiced[23]. Perhaps, therefore, when the task is novel or difficult, the focus of attention adjusts and zooms in to capture less than four chunks, so as to leave more attention free to carry out the task itself. With practice, the task becomes more automatic and attention can be used to hold more chunks at once. A slightly different suggestion[24] is that the focus of attention itself only holds one chunk, but there is a mental region associated with that focus that holds up to four chunks.

Some procedures are highly controversial. Let us return to the enumeration procedure discussed by Miller[1]. Subsequent work has set the limit for rapid enumeration without counting, called *subitizing*, not at seven but at about four objects[25]. Some have suggested that subitizing has nothing to do with a limit in primary memory capacity, but rather with the observation that spatial patterns can be more easily recognized when they consist of fewer objects because, as the number of objects in the display increases, the number of distinguishable patterns skyrockets[26]. This might explain why primitive skills of enumeration of small numbers exist even in infants and non-human animals[27]. However, some research argues against that interpretation. It has been found that elderly individuals cannot subitize as many objects as young adults can[28], yet there is no reason to suspect that the elderly lose the ability to detect known patterns; a great deal of previous research does suggest, though, that their primary-memory capacity is diminished relative to young adults.

Has One Legend Been Replaced By Another?

In this chapter, we began by discussing a simple answer to the question of what primary memory capacity is: that primary memory can hold seven chunks or meaningful units. This answer was shown to have some basis in the facts, but overall it was shown not to be a general rule, and therefore was said to be a legend.

However, it should be said that simple answers are not, in principle, bad. One of the goals of science is to find simple rules to explain the available evidence in a comprehensible manner. What makes the simple rules unacceptable is just when they are shown not to match the facts. By analogy, in the realm of physics, it was not a bad move for Isaac Newton to propose a simple law of gravitational force, because it helped explain the data of planetary motion collected by Tycho Brahe and the regularities of planetary motion derived from the data by Johannes Kepler. The laws of gravity could be clearly observed only in situations in which wind resistance was eliminated or taken into account; just as, we have suggested, the capacity of primary memory can be clearly observed only in situations in which rehearsal and chunking have been eliminated or taken into account.

Just as the more comprehensive understanding of gravity by Albert Einstein eventually displaced the simpler gravitational law of Newton, a more comprehensive understanding of primary memory capacity is bound to come along and replace the simple generalization[12, 13] that people can remember on average three or four chunks of information. Until that time, however, the limit of three or four serves as a useful guideline for research and theory, as did the gravitational constant for many years. What is likely to advance us to the next level, beyond a new legend of three or four, is a better understanding of the long-term memory processes involved in chunking, a topic emphasized in the seminal work that launched the modern research on primary memory: the article published by George Miller[1] in 1956, about 50 years before the present chapter went to print.

References

1. 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.
2. James, W. (1890). *The principles of psychology*. NY: Henry Holt.
3. Miller, G.A. (1989). George A. Miller. In Lindzey Gardner, ed., *A history of psychology in autobiography*, Vol. VIII. Stanford, CA: Stanford University Press. (pp. 391-418)
4. Ericsson, K.A., Chase, W.G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*, 1181-1182.
5. Rouder, J.N., Morey, R.D., Cowan, N., & Pfaltz, M. (in press). Learning in a unidimensional absolute identification task. *Psychonomic Bulletin & Review*.
6. Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74* (Whole No. 498.)
7. Chase, W., & Simon, H.A. (1973). The mind’s eye in chess. In W.G. Chase (ed.), *Visual information processing*. New York: Academic Press. (pp. 215-281)
8. Gobet, F., & Simon, H.A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory, 6,* 225-255.
9. 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.
10. Johnson, M.F. (1969). The role of chunking and organization in the process of recall. In G.H. Bower & J.T. Spence (eds.), *Psychology of learning and motivation. Vol 4*. Oxford, England: Academic Press. (pp. 171 - 247).

11. 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.

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

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

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

15. 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.

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

17. 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.

18. 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.

19. Jefferies, E., Lambon Ralph, M.A., & Baddeley, A.D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working

memory. *Journal of Memory and Language, 51*, 623-643.

20. Vogel, E.K., & Machizawa, M.G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*, 749-751.
21. Todd, J.J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*, 751-754.
22. McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 817-835.
23. Verhaeghen, P., Cerella, J., & Basak, C. (in press). A Working-memory workout: How to expand the focus of serial attention from one to four items, in ten hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*
24. 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.
25. Mandler, G., & Shebo, B.J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111*, 1-22.
26. Logan, G.D., & Zbrodoff, N.J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review, 10,* 676-682.
27. Gallistel, C.R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences, 4,* 59-65.
28. 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.

Figure Captions

Figure 1. Illustration of a procedure used by Cowan, Chen and Rouder[15] to examine the capacity limit of serial recall expressed in chunks.

Figure 2. Results of a study by Cowan, Chen, and Rouder[15] of the information produced in the serial recall of 8-word lists. The results are averaged over two experiments. The blue triangles show that the average number of words recalled in the correct serial positions increased as a function of the amount of training with pairs of words. (A similar trend was observed for words recalled regardless of the serial positions.) The red circles show that the average number of chunks that were recalled nevertheless remained constant across these training conditions. Chunks included words recalled as singletons, and also pairs of words that were presented together within the list and recalled together with the pair intact.

Figure 3. Illustration of the array-comparison procedure of Luck and Vogel[16] as adapted by Morey and Cowan[17].

Figure 4. Results of a study of the effect of a verbal memory load on the retention of an array of colored squares to be compared with a second array[17]. A formula (Cowan[13], p. 166) was used to estimate visual memory capacity expressed as the number of squares retained, which was then averaged across arrays with 4, 6, or 8 squares. The key finding is that although repeating 7-digit load had a strong effect, especially when the load was repeated incorrectly, repeating a known 7-digit number (the participant's own telephone number) had little effect. Therefore, it was the demand on attention rather than articulation per se that disrupted retention of the array of colored squares.

Figure 1

Training Conditions (including 4 presentations of each word)

1. Words presented 4 times as singletons, but never paired (0-pairing)
2. Words presented 3 times as singletons, and 1 time paired (1-pairing)
3. Words presented 2 times as singletons, and 2 times paired (2-pairing)
4. Words never presented as singletons, but 4 times paired (4-pairing)

Presentations were randomly mixed as in *box, hat, dog-shoe, box, girl, desk, tree-brick, hat, dog-shoe...*

Serial Recall Test

For each training condition, a list of 4 pairs of words was presented using known pairs.
Pairs from a condition were randomly arranged in a list, as in *tree-brick, dog-shoe, man-tank, rock-coin*.

Cued Recall Test (before or after the serial recall test)

The first word in a pair was presented and the correct response was the second word, as in *dog - ???*

Figure 2

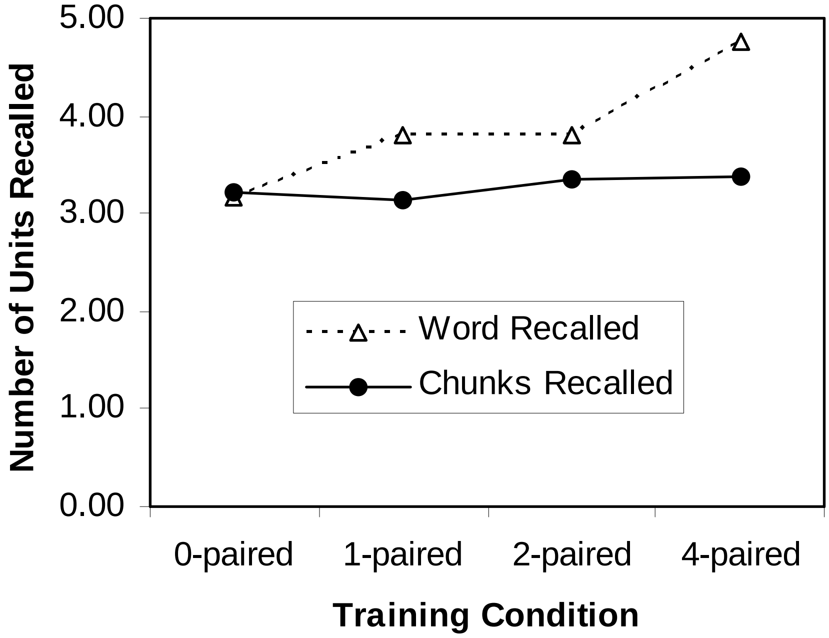

5.00
4.00
3.00
2.00
1.00
0.00
Number of Units Recalled
Word Recalled
Chunks Recalled
0-paired
1-paired
2-paired
4-paired
Training Condition

Figure 3

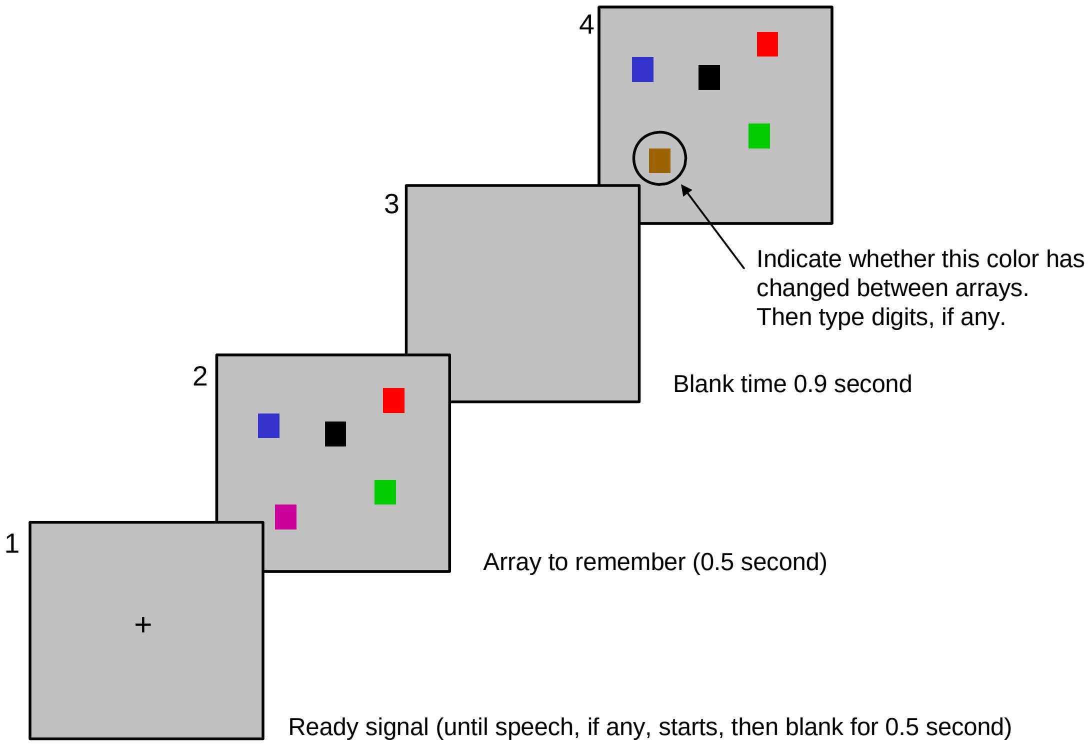
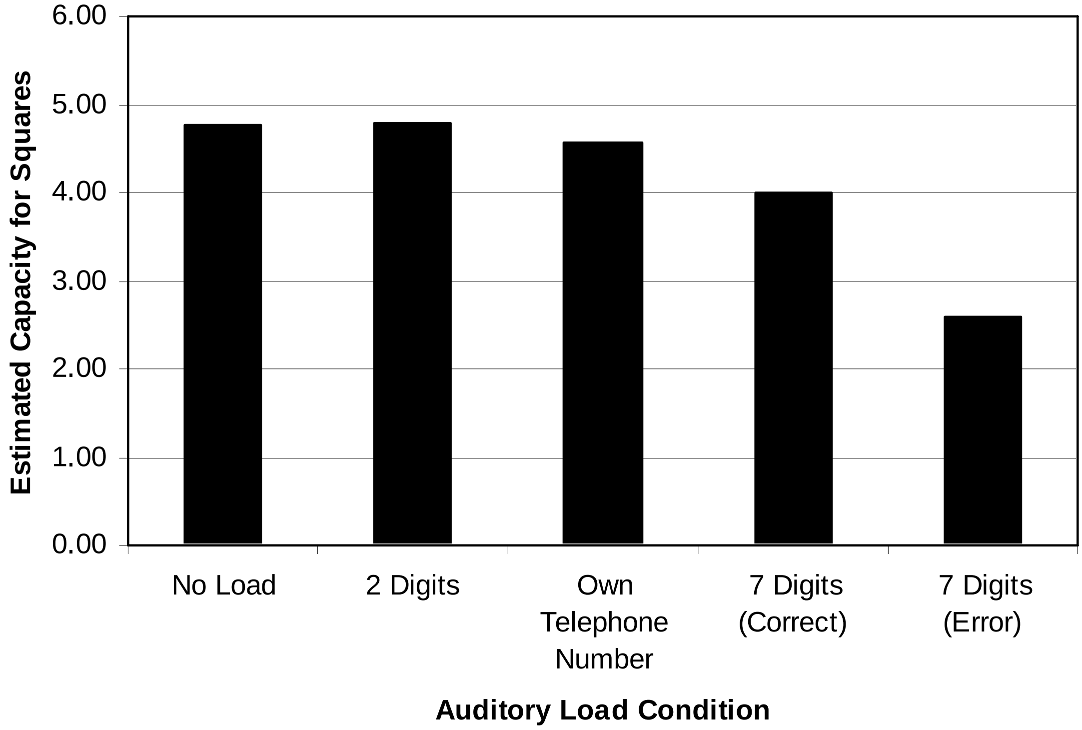

4
3
Indicate whether this color has
changed between arrays.
Then type digits, if any.
2
Blank time 0.9 second
1
+
Array to remember (0.5 second)
Ready signal (until speech, if any, starts, then blank for 0.5 second)

Figure 4


6.00
5.00
4.00
3.00
2.00
1.00
0.00
Estimated Capacity for Squares
No Load
2 Digits
Own Telephone Number
7 Digits (Correct)
7 Digits (Error)
Auditory Load Condition