Visual and auditory working memory capacity

Nelson Cowan

A considerable amount of cognitive-behavioral research has been conducted on working memory. Definitions vary, but a theory-independent definition might state that working memory is the collection of mental processes that permit information to be held temporarily in an accessible state, in the service of some mental task. The nature of the task can vary widely and can include immediate recall, reading or listening comprehension, reasoning, or problem-solving. In listening comprehension, for example, it is often the case that the intended meaning of a word within a sentence is unclear until subsequent words in the sentence are presented. It is necessary to hold words in mind in some form until their meanings can be interpreted in light of the remainder of the sentence. In reasoning, assumptions and facts must be held in mind and considered together until conclusions can be deduced from them. It has been clear that the capacity of working memory is limited ever since George Miller described various research studies suggesting that people can recall at most about seven independent, meaningful items or ‘chunks’ at a time.

Although limits to working memory are easily observed, it is much more difficult to determine what specific mental faculties underlie the observed limits. For example, the capacity of the phonological memory, lists of words is better when the lists contain inter-syllabic pauses, not relatively quickly, and this ‘word-length effect’ occurs to some extent even when the short-word and long-word sets actually comprise the same words, but with instructions to pronounce them quickly versus slowly. To account for stimulus-dependent working memory limits as well as age differences in working memory, various researchers have proposed working memory systems that include multiple component processes. Baddeley proposed a system that includes a ‘central executive’ process that makes use of a passive, time-limited phonological store along with a ‘covert rehearsal’ process for verbal recall, and a passive, time-limited ‘visuospatial store’, possibly with another covert rehearsal process, for visual recall. In my own theoretical writings I have proposed that working memory is composed of a capacity-limited focus of attention, along with a temporarily activated portion of the information in permanent memory, which extends beyond the focus of attention to include some automatically activated information (see Fig. 1). When researchers use the term ‘working memory’, some of them seem to be referring to all of the temporarily activated information. It is also likely that certain inactive portions of memory can be stored in a way that allows them to be recalled (or reactivated) quickly. For example, in a reasoning problem involving rainbow colors, encoding the seven colors of the rainbow as the name ‘Roy G. Biv’ and keeping that name in mind makes the color names easily accessible while using up perhaps only one to three items of working memory capacity.

There have been some attempts to go beyond the observed working memory limits to glean the limits of the underlying processing components. Broadbent suggested, on the basis of past evidence, that the true capacity limit is about three items (presumably when the contributions of semantically useful processes such as rehearsal and long-term memorization have been eliminated). For example, this is about the number of items that can be recalled without error across many trials, and about the maximum number of items that can be grouped together into a single ‘chunk’, although the actual limit may be closer to four. Other researchers have proposed similar capacity limits of three or four items in the number of processing channels for visual search, the number of items that can be enumerated quickly, without a slow, serial counting process, and the number of moving visual items that can be tracked at the same time. A similar limit of about four items has been found when subjects encounter a spatial array of printed characters or a spatiotemporal array of spoken characters and must report them all. It is not clear if all of these similar limits are related; if so, perhaps they reflect the capacity limit of the focus of attention.

Recently, Luck and Vogel have contributed to this area in an important way. They first presented a spatial array of colored squares or rectangles on every trial. The second presentation was another array that could differ from the first array in the color of one item. Subjects were able to carry out the task well only if the first array, the one to be held in memory, contained four or fewer items. (The same pattern was obtained in an experiment in which a single item within the second array, the one that sometimes differed from the first array, was marked with a small rounding square in order to limit the decision to that one item.) These results extend the previously observed capacity limit to nonverbal visual stimuli, and to a situation in which there was only one decision to be made (in the case when there was a single probe item). A few other results warrant special mention and discussion.

First, the observed capacity of visual working memory was not reduced when subjects had to hold in mind two digits during a visual memory trial, to be recalled immediately afterwards. One might expect a reduction of visual working memory if both verbal and visual representations were held in the same capacity-limited store. However, it is possible that the two verbal items could be held entirely in the form of a passive phonological store and rehearsal process without taking up space in the capacity-limited store or focus of attention. If the verbal memory load were increased further or accompanied by a rehearsal-blocking task, it might well be shown to reduce the observed capacity of visual working memory.

Second, the capacity limit was found to be the same (about three items) no matter whether the discrepancy between displays occurred in one feature, two features or four features of each object. Thus, the capacity is apparently for the integration of all features per se. This at first may seem curious, in view of the fact that other research suggests that links between features must be perceived one object at a time rather than all objects in parallel. Possibly, subjects read the items into working memory one at a time but are still limited to about four such items on a trial. If this is true, despite the short durations of the arrays (>100 ms), items must remain in a post-stimulus sensory memory, and, therefore, memory should be greatly curtailed by a pattern mask immediately following the first array in a trial.

For arrays larger than four it is not even absolutely clear whether subjects encode a single feature of each item in a visual array (either serially or in parallel) or do a partial encoding of all of the items (e.g. about half the features of each item in an eight-item array). In
other words, the basis of the four-item limit is still unclear. Various research strategies could be of use here. For example, an item-by-item analysis could theoretically reveal that it is usually the four items closest to the fixation point that are encoded. If, instead, all items are partially encoded, then it would be possible to improve performance by changing more than one feature of the target object between presentations in the same trial, increasing the chances that at least one of the critical features had been encoded by the subject. Working memory has also been a popular topic within recent neuromaging studies. It is important to realize that there is still considerable behavioral work to be done before it will become clear what the behaviors are that might be explained through brain processes.

References
2 Miller, G.A. (1956) The magical number seven, plus or minus two: Some limits on our capacity to process information Psychol. Rev. 62, 81-97
4 Cowan, N. et al. (1987) There are two word-length limits in the human short-term memory: opposed effects of duration and complexity Psychol. Sci. 2, 206-208