The behavioral results reviewed by Cowan and our fMRI results together indicate that verbal STM loads below a "4 ± 1" capacity limit invoke a specific STM system mediated in part by left inferior PFC. Supracapacity verbal STM loads invoke multiple additional executive and attentional systems mediated by bilateral dorsolateral prefrontal and cingulate cortices. STM circuitry extends beyond these areas to at least parietal and cerebellar regions, but these prefrontal findings offer independent, convergent evidence for Cowan's persuasive argument.

Characterizing chunks in visual short-term memory: Not more than one feature per dimension?

Werner X. Schneider, Heiner Deubel, and Maria-Barbara Wesenick

Institute of Psychology, Ludwig-Maximilians-University Munich, D-80802 Munich, Germany. **{wxs; deubel; wesenick}@psy.uni-muenchen.de**
www.paed.uni-muenchen.de/mip/psych/deubel/wwwdocs/index.htm

Abstract: Cowan defines a chunk as "a collection of concepts that have strong associations to one another and much weaker associations to other chunks currently in use." This definition does not impose any constraints on the nature and number of elements that can be bound into a chunk. We present an experiment to demonstrate that such limitations exist for visual short-term memory, and that their analysis may lead to important insights into properties of visual memory.

To determine the capacity limit of short-term memory (STM) can be a tricky business with a number of potential pitfalls. Cowan provides a careful and much needed analysis of these pitfalls that can confound estimates of the real memory capacity with other limitations in the system. After reviewing a large number of studies that try to avoid these problems, Cowan concludes that the limit of (STM) is about 4 chunks, rather than the classical 7 ± 2. Concerning visual working memory, this capacity estimate has indeed been put forward by several authors for quite some time. Based on experimental evidence from studies on transsaccadic memory (e.g., Irwin 1992) and on visual STM (e.g., Shibuya & Bundesen 1988), one of us reached a similar conclusion in a theoretical analysis of visual working memory (Schneider 1999).

However, knowing how many chunks can be retained is by no way sufficient if we want to know how much information can be stored in STM – to answer this question, it is absolutely necessary to know more precisely how a visual chunk can be characterized. Indeed, we think that a central and important issue of future research will be to analyse the limits of STM in terms of the mechanisms by which more basic elements are formed into chunks, and also to study how attention and specific tasks determine this chunking.

Imagine a scene consisting of four "objects," each being a square made up of a 3*3 raster of differently coloured parts. An intuitively obvious prediction would be that these four objects *cannot* be retained as well as four homogeneously coloured squares. Amazingly, the data from one of Luck and Vogel's (1997) experiments seemed to suggest otherwise. In this experiment, subjects had to retain such stimuli as shown in Figure 1 that were defined by various colour-colour-conjunctions. The results seemed to show that objects that are made of a conjunction of two colours could be retained equally well as objects with just one colour value per object. This implies that visual chunks may contain at least two feature values for the colour dimension.

For storage of visual information such an analysis is possibly somewhat easier than for verbal material, which is the focus of Cowan's article. In the visual format it is intuitively plausible to assume that chunks are direct reflections of consciously perceived visual "objects." A frequent assumption in vision is that an object can be described in terms of its basic visual dimensions, such as its colour, shape, or motion. Within each dimension, the stimulus can be specified by "features" or "feature values" (e.g., Treisman &

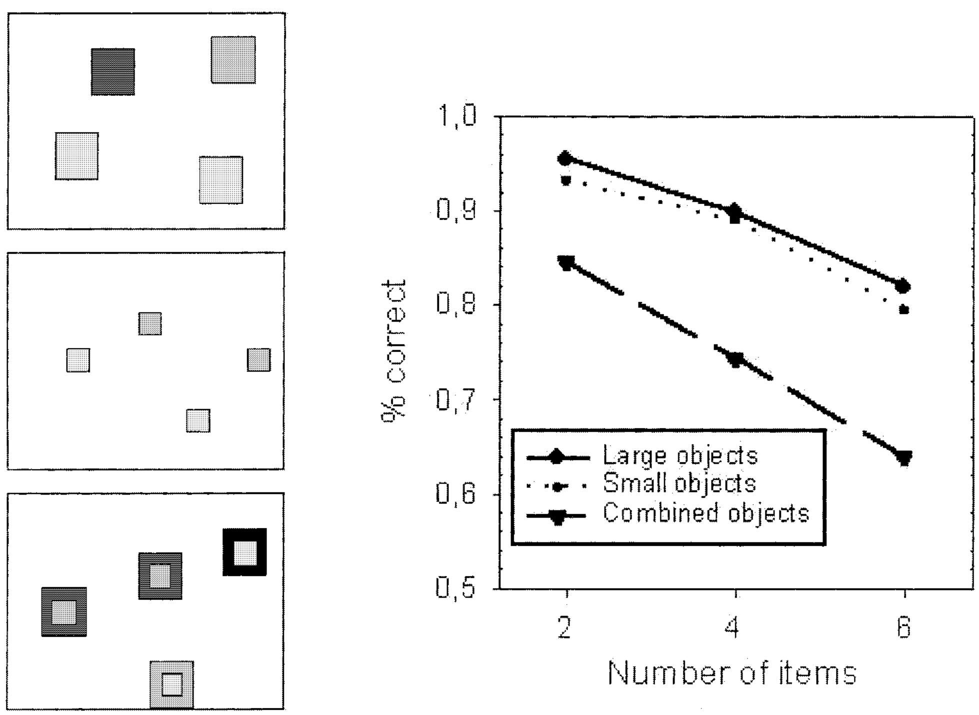

Figure 1 (Schneider et al.). Left: Stimuli used in the experiment. Subjects had to memorize configurations of 2, 4, or 6 coloured squares for 900 msec. After the retention period, either the same display or a display with one of the squares differing in colour reappeared; subjects were asked whether they had detected a change. Squares could be large (0.75*0.75 deg) or small (0.35*0.35 deg), or conjunctions of both. Right: Percent correct report as a function of the number of items presented.

Gormican 1988) For instance, the dimension "colour" of an object can be characterized by one of the features "red," "blue," "purple," and so on. After breaking down a visual object into these more basic elements, the question arises as to whether the above definitions of "dimensions" and "features" in vision constitute an appropriate basis for characterizing chunks in visual memory. In other words, one may ask (1) whether chunking in visual working indeed functions by binding features into integrated visual objects, and (2) analyse the rules and limitations of this binding.

With regard to the first question there is recent evidence from Luck and Vogel (1997, also referred to in sect. 3.1.1), suggesting that the capacity limit of visual STM indeed refers to feature bundles in the form of objects. In their experiments, subjects were required to retain simple geometrical visual objects made up of feature conjunctions such as of a certain colour, orientation, and length. The data showed that objects defined by conjunctions of two or more dimensions (e.g., a line of a certain orientation, colour, and length) can be retained as well as objects defined by only a single dimension (e.g., orientation only). For any of these combinations, the estimated memory capacity was about four objects.

The second question is strongly related to the problem of how many of these basic elements can be bound into a single chunk, and whether there exist limitations as to the possible combination of features.

In our own experiments, we attempted to replicate this surprising finding (Deubel et al., in preparation). We used the same stimuli as Luck and Vogel (1997), and identical experimental parameters such as presentation and retention times. Our experimental results (Fig. 1) clearly show that retention of objects defined by a conjunction of two colours leads to a strong drop in performance, as compared to the condition in which the objects consisted of one colour only. So, external objects with two colours seem to require two chunks for the internal coding. This finding is in obvious contrast to the result of Luck and Vogel (1997). The reason why we could not replicate their data is unclear to us, however, in an independent study, Wheeler and Treisman (submitted) recently reported a finding similar to ours.

These data are clear evidence that there exist prominent limitations to chunking in visual memory. As a possible, preliminary rule of thumb suggested by the result, one may assume that a visual chunk can consist of not more than one feature per dimension, that is, one colour, one shape primitive, and so on. A further, yet unresolved important issue in this context is the question whether there is also a limit in the number of possible dimensions that define a chunk. Luck and Vogel (1997) found no limit (i.e., no drop in memory performance) up to a conjunction of four different dimensions (colour, orientation, length, gap). However, it might be that a limitation larger than that can indeed be found.

The empirical task of the future will be to determine more precisely the limitations of chunking and how they relate to visual features and dimensions. Indeed, we think that the paradigm presented here offers a promising experimental approach to answer questions about the nature of chunks in vision. Measuring memory performance for a variety of stimuli and features could reveal the basic dimensions and features in vision in a very straightforward way: If adding the feature in question to the stimuli leaves the memory capacity (in terms of number of objects) unaffected, one may conclude that it is really a basic visual feature, forming an elementary part of a visual chunk.

Cowan defines a chunk as "a collection of concepts that have strong associations to one another and much weaker associations to other chunks currently in use." This definition does not impose any constraints on the nature and number of elements that can be bound into a chunk. Our experiment is a demonstration that such limitations exist, and that their analysis may lead to important insights into properties of visual memory.

The magical number 4 in vision

Brian J. Scholl[a] and Yaoda Xu[b]

[a]*Department of Psychology, Vision Sciences Laboratory, Harvard University, Cambridge, MA 02138;* [b]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139.*
scholl@wjh.harvard.edu **yaoda@psyche.mit.edu**
www.wjh.harvard.edu/~scholl

Abstract: Some of the evidence for a "magical number 4" has come from the study of visual cognition, and Cowan reinterprets such evidence in terms of a single general limit on memory and attention. We evaluate this evidence, including some studies not mentioned by Cowan, and argue that limitations in visual processing are distinct from those involved in other memory phenomena.

Cowan's discussion of the "magical number 4" synthesizes evidence from domains which are rarely discussed together. In particular, Cowan draws on work from the study of visual cognition – such as studies of subitizing (sect. 3.3.2) and multiple object tracking (sect. 3.3.3) and attempts to reinterpret such evidence in terms of a general memory limitation, which he suggests is a reflection of the underlying capacity of the "attentional focus" (a thesis which is discussed in Cowan 1995, but which he does not argue for in his target article). Here we note additional evidence for a limit of approximately 4 objects in certain types of visual processing, and discuss why these limits are probably distinct from those involved in other (e.g., verbal) tasks.

Additional evidence from visual cognition. Additional evidence for a "magical number 4" in visual processing comes from studies of infants, normal adults, and neuropsychological syndromes. Recent looking-time studies with infants have suggested that they are able to keep track of arrays of objects through additions and subtractions, but only if there are less than 4 objects in these arrays (e.g., Uller et al. 1999; Wynn 1992), and this evidence has been interpreted in terms of developing mechanisms of visual attention (e.g., Carey & Xu, in press; Scholl & Leslie 1999). In normal adults, there appears to be a limit of 4 on the number of objects which can receive prioritized processing due to attentional capture (Yantis & Johnson 1990), and the number of items which can be simultaneously examined in a visual search for a change (Rensink 2000).

Finally, it has been shown that bilateral lesions of the parietal lobes in Balint's syndrome can reduce visual processing capacity. Patients with Balint's syndrome have great deficits in perceiving complex visual scenes, although their ability to recognize individual objects is usually preserved (for a review, see Rafal 1997). Dehaene and Cohen (1994) studied visual enumeration in 5 Balint's patients and found that these patients could enumerate sets of 1, 2, and sometimes 3 items correctly, but not sets comprising more than 3 items. Reaction time slopes for these patients were flat for set sizes of 1 and 2 items, but increased sharply for set sizes of 3 or more items. Treisman and colleagues (Friedman-Hill et al. 1995; Robertson et al. 1997) reported another Balint's patient who could not correctly enumerate more than one or two objects even when he was aware that more were present. In rare and extreme cases, Balint's patients report seeing only one object when presented with multiple objects (e.g., Coslett & Saffran 1991).

Specific visual limits or general memory/attention limits? Cowan views such evidence as continuous with data concerning the number of chunks which can be simultaneously active in short term memory (STM). In contrast, we think there are good reasons to resist this reinterpretation, and to view the limits on visual processing as separate from those involving verbal and other non visual material. (In this respect we take a position similar to that of Miller 1956 who suspected that STM limits and subitizing limits were independent.) Given space restrictions, we will largely restrict our discussion of this issue to the evidence which Cowan does discuss in his target article: subitizing (wherein observers can determine the cardinality of sets with less than 5 items roughly in parallel and without errors) and multiple object tracking (MOT;