This commentary appears as:

Scholl, B. J., & Xu, Y. (2001). The magical number 4 in vision. *Behavioral and Brain Sciences*, *24(1)*, 145 - 146.

The magical number 4 in vision

Brian J. Scholl[1] and Yaoda Xu[2]

1 Vision Sciences Laboratory, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA, 02138, USA. **scholl@wjh.harvard.edu, http://wjh.harvard.edu/~scholl**

2 Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, 02139. **yaoda@psyche.mit.edu**

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 (section 3.3.2) and multiple object tracking (section 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. Wynn, 1992; Uller et al., 1999), and this evidence has been interpreted in terms of developing mechanisms of visual attention (e.g. Scholl & Leslie, 1999; Carey & Xu, to appear). 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; wherein observers can attentionally track up to 4-5 independently and unpredictably moving identical items in a field of identical distractors).

Cowan presents only a few arguments for interpreting these phenomena in terms continuous with general STM limits. For MOT he provides no arguments, simply stating that one could use a general STM-based theory to explain performance. (Such an explanation, it seems to us, could not easily account for the strong dependency of MOT performance on subtle visual details such as the type of accretion and deletion behind occluders; Scholl & Pylyshyn, 1999.) For subitizing, he notes the vision-based theory of Trick and Pylyshyn (1994), and argues against it mainly by appeal to two phenomena. First, he suggests that the 'pop-out' alluded to by Trick and Pylyshyn can also occur for larger numbers of items, for example "when all of the eggs [in a carton] pop out against the surrounding carton" (Section 3.3.2). This, however, is clearly not the type of pop-out that Trick and Pylyshyn (and others who have investigated visual search) have in mind, since the eggs in this case do not pop out as individuals, but as a group. Second, Cowan suggests that focused central attention is more important to enumeration than is suggested by Trick and Pylyshyn's theory, since other researchers (Atkinson et al., 1976; Simon & Vaishnavi, 1996) studying the enumeration of dots in afterimages have claimed that observers cannot enumerate sets greater than 4 without eye-movements. This claim is false, however, and the limits these investigators found were due to the confounding effects of crowding (He et al., 1997).

Beyond Cowan's arguments, we think there are several additional reasons to view these limits as distinct from those involved in verbal STM. First, viewing them as identical seems to necessitate a prediction that one should not be able to track 4 targets in the MOT task and simultaneously acquire and hold 4 verbally-presented items in STM. However, this is trivial to do, and such tasks seem not to interfere at all. (In an informal test, two observers tracked 4 in 8 items for 10 s with an accuracy of 87.5% averaged over 10 trials. When they also had to remember 4 random digits presented auditorily as the targets were being specified, they tracked with an accuracy of 92.5%, and made no errors on the memory task.) Cowan notes in section 4.2 of the target article that such evidence against a single capacity limit could be explained away by appeal to attentional switching back and forth between the two tasks, but in this respect MOT is an ideal foil, since one can succeed in the task only by continuous tracking (Pylyshyn & Storm, 1988).

Second, an explanation based on a single general limitation of memory or attention predicts that these limits should stand or fall together in neuropsychological impairments, which they do not. For example, none of the Balint's patients mentioned above exhibited deficits in short-term memory span. There are patients who, after lesions in the left hemisphere language areas, exhibited reduced STM span despite normal speech production in some cases (e.g., Baddeley, 1986; Shallice & Warrington, 1970). However, none of these patients showed any signs of Balint's symptoms or deficits in visual processing. Moreover, although these patients showed very poor retention of auditorily presented digits, with a span in the region of two items, they usually showed better retention of visually presented digits, with a span in the region of 4 or 5. These double dissociations in lesion sites and patient performance argue strongly against the notion that a common capacity limitation underlies capacity limited performance in both verbal and visual tasks.

Visual objects vs. chunks in memory. The view that these limitations in visual processing are distinct from those involved in other memory phenomena is further strengthened by the fact that the 'units' of processing in each case are quite different. The 'chunks' of memory can be almost infinitely flexible in their composition, and are thus defined by Cowan and others simply in terms of association networks (see section 1.3). This flexibility is in marked contrast to the units of visual attention — visual objects — which appear to be characterized by highly constrained and inflexible rules (Scholl, to appear). In MOT, for instance, observers can track 4 dots in a field of 8 dots, but completely fail when trying to track 4 line endpoints in a field of 4 lines (and thus 8 endpoints). In general, very specific rules involving connectedness and part-structure seem to determine whether a feature cluster can be tracked in MOT (Scholl et al., submitted). Similarly, in visual short-term memory studies using a change detection paradigm, color and orientation features are best remembered if they belong to the same part of an object and less well remembered if they belong to different parts of an object (Xu, submitted). All of these constraints are in marked contrast to the robustness and flexibility of potential STM chunks with verbal materials.

We think the considerations discussed here provide good reasons for thinking that the limits of approximately 4 involved in various types of visual processing are distinct from other similar STM limits. We remain agnostic on the question of why there should exist similar independent limits. It could be for the teleological and computational reasons discussed by Cowan (in section 4.1), or it could be — as George Miller (1956) suspected of the similarity of memory capacity and subitizing limitations — "nothing more than a coincidence."

References

Atkinson, J., Campbell, F., & Francis, M. (1976). The magic number 4 +/- 0: A new look at visual numerosity. *Perception, 5*, 335 - 342.

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

Carey, S., & Xu, F. (to appear). Infant knowledge of objects: Beyond object files and object tracking. *Cognition*, Special Issue on 'Objects and attention'.

Coslett, H. B., & Saffran, E. (1991). Simultanagnosia: To see but not two see. *Brain, 114*, 1523-1545.

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

Dehaene, S., Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnosic patients. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 958-975.

Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995). Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science, 269*, 853-855.

He, S., Cavanagh, P., & Intriligator, J. (1997). Many objects can be enumerated in afterimages when they are not crowded [Abstract]. *Abstracts of the Psychonomic Society, 2*, 3.

Miller, G. A. (1956). The magical number seven, plus of minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 176 - 185.

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

Rafal, R. D. (1997). Balint syndrome. In T. Feinberg & M. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 337 - 356). New York: McGraw-Hill.

Rensink, R. A. (2000). Visual search for change: a probe into the nature of attentional processing. *Visual Cognition, 7*, 345 - 376.

Robertson, L., Treisman, A., Friedman-Hill, S., & Grabowecky, M. (1997). The interaction of spatial and object pathways: Evidence from Balint's syndrome. *Journal of Cognitive Neuroscience, 9*, 295-317.

Scholl, B. J. (to appear). Objects and attention: The state of the art. *Cognition*, Special Issue on 'Objects and attention'.

Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In E. Lepore & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 26 - 73). Oxford: Blackwell.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38*, 259 - 290.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (submitted). What is a visual object? Evidence from multiple object tracking. Manuscript submitted for publication.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22,* 261-273.

Simon, T., & Vaishnavi, S. (1996). Subitizing and counting depend on different attentional mechanisms: Evidence from visual enumeration in afterimages. *Perception & Psychophysics, 58*, 915 - 926.

Trick, L., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80 - 102.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development, 14*, 1 - 43.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749 - 750.

Yantis, S., & Johnson, D. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 812 - 825.

Xu, Y. (submitted). Integrating color and shape in visual short-term memory for objects with parts. Manuscript submitted for publication.