ATTENTION: EXPORTING VISION TO THE MIND

Patrick CAVANAGH
Department of Psychology, Harvard University
33 Kirkland Street, Cambridge, MA 02138, USA
patrick@wjh.harvard.edu

Attention is the gateway to visual awareness and it imposes the final limit on what we can consciously experience in our visual world. We have evaluated the grain of the spatial sampling available to attention and discovered that the smallest regions which can be isolated by attention are surprisingly coarse. The properties of attentional resolution suggest that the locus of attentional selection is at a stage beyond primary visual cortex. Based on our results, I suggest that attention is not the final stage but just one direction taken in the analysis of visual information, a direction which leads specifically to a sparse conscious model of the world. In particular, a central role of attention is to translate descriptions generated by the primary visual system into a highly constrained public code that is broadcast to other modules of the brain.

INTRODUCTION

We normally pay attention to a visual event, say the appearance of a celebrity within our field of view, by directing our eyes so that the area of interest falls at the center of gaze. However, we can also increase our processing capacity by deploying attention to the event, without moving the eyes. One of the most enduring metaphors for attention is that of a "spotlight" (Eriksen & Hoffman, 1972) illuminating a small region of visual space where details are more easily read out -- to the detriment of surrounding areas. However, a number of experiments have shown that the attention can be directed to noncontiguous locations (Castiello & Umilta, 1992; Kramer & Hahn, 1995) and even used to track as many as four or five randomly moving items simultaneously (Pylyshyn & Storm, 1988). Based on work from our laboratory, we suggest that attention is more like an image than a single spotlight and that it has image-like properties of resolution.

In general, the selective filter of attention is often taken to be the last step in perception, feeding integrated representations into the final model of the world which fills our visual awareness. I would like to suggest a different view, one that reorganizes the elements of vision, attention, and conscious perception in a different structure. Many have proposed that there are two or more parallel streams of visual analysis (DeYoe & van Essen, 1988; Goodale & Milner, 1992; Ungerleider & Mishkin, 1982). In line with these suggestions, it is likely that attention is not the final bottleneck of all visual analysis, but actually affects just one direction taken by visual information, a direction which leads specifically to conscious perception, a condensed product for export to the rest of the mind. There may be several important paths taken by visual information and the content of conscious perception should not be taken as the only indicator of the limits and properties of visual representation.

THE GRAIN OF ATTENTION

Figure 1 demonstrates the basic point of the spatial resolution of attention. While fixating the central dot, it is easy to visually resolve the lines to the left -- they are numerous, black, thin, and vertical. It is much harder, however, to attend to individual lines, to count them or step through them one at a time (without moving your eyes from the central dot). Visual resolution is conventionally measured as the finest sinusoidal grating that can be seen when presented at maximum (100%) contrast (Campbell & Gubisch, 1966). A observer can normally resolve up to 55 cycles per degree. We can define attentional resolution in a similar manner. What is the finest spacing of the bars of a grating at maximum contrast that allows the observer to index each bar. In tracking and counting tasks, we find that the limit of attentional resolution is about one order of magnitude coarser than that of visual resolution.

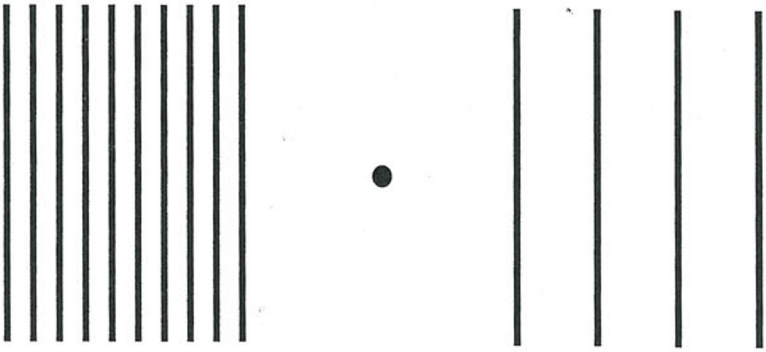

Figure 1. Fixate the central dot and attempt to pay attention to each line in turn. A simple way to do this is to try to count them or step through them one at a time. On the left, it is very difficult to isolate individual bars. We say that the density of this display exceeds the spatial resolution limit of attention. On the right, however, it is easier and so we claim that this display is within the resolution limits of attention.

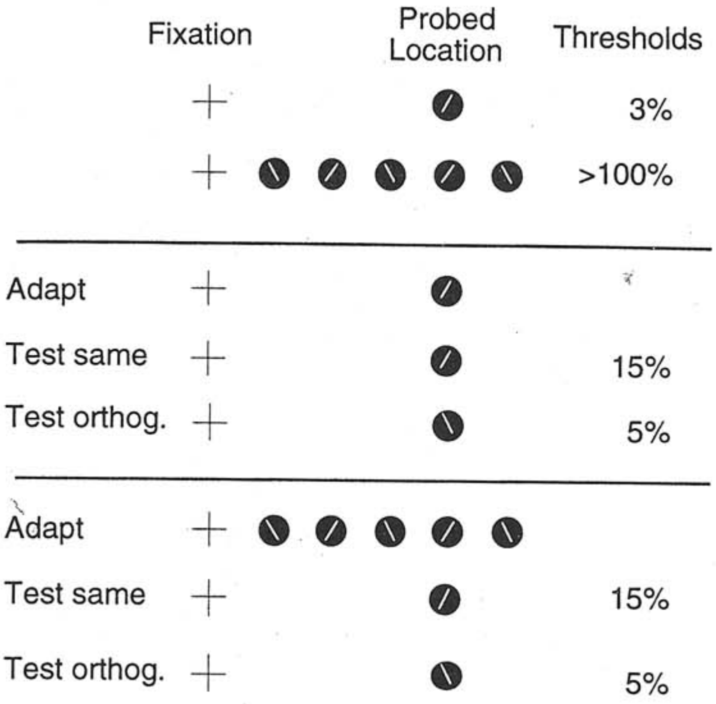

Figure 2. The first panel shows the contrast of the grating (shown here as a single bar) at the probed location required for its orientation to be judged with 75% accuracy. In the crowded case, responses were basically at chance levels. The orientation information was inaccessible to conscious report. Following adaptation, we measured the contrast required to report the orientation of a single grating at the location of the adaptation. Three times more contrast was required if the test had the same orientation as the adapting grating. The same orientation specific effect was found whether the observer was aware of the orientation (single adapt, middle panel) or not (crowded adaptation, bottom panel). Adapted from He, Cavanagh, and Intriligator (1996) but depicted for simplicity as single bars within a circular disk arrayed horizontally. In the experiment the entire disk was filled with a grating of uniformly spaced bars and the disks were aligned vertically above fixation.

CROWDING

In the grating above, the cost of not being able to individuate the bars is relatively small -- we do not know how many there are but we know they are all vertical and thin. If each bar were subtly different, however, we might not be able to report the features of individual bars. In severe crowding like that on the left in Figure 1, a target embedded in a dense array of distractors can not be selected by attention, hence can not be consciously scrutinized, independently of the distractors. We claim that this crowding effect is due to insufficient spatial attentional resolution and this claim is further supported by the next experiment.

To measure the consequences of crowding, we (He, Cavanagh, & Intriligator, 1996) used an effect of adaptation which we know to be mediated by structures in the visual system located in primary visual cortex or beyond. Specifically, adaptation to a grating in a particular orientation will raise the threshold for subsequently detecting a grating in that same orientation much more than detecting a grating in the orthogonal orientation. This orientation selective adaptation has been linked to neurons at the primary visual cortex and beyond (Movshon & Lennie, 1979). In our study, we adapted our observers either to a single grating, or a crowded grating array. In this latter, crowded case, the adapting grating at the target position was unavailable to conscious perception (see Figure 2, top), and identification of the orientation of the grating was at chance levels. However, we found the same magnitude of orientation-selective adaptation effect in both the single and crowded adaptation conditions (Figure 2, middle and bottom).

A grating that is unavailable to our conscious perception (due to crowding) is nevertheless as powerful as a single, clearly visible grating in producing orientation-specific adaptation effects. This suggests that crowding must happen at a stage beyond the site of adaptation which itself must be at least at primary visual cortex (no orientation analysis occurs at earlier levels). This result is consistent with the proposal that activation of neurons in primary visual cortex alone is not sufficient for visual awareness (Crick & Koch, 1995; Koch & Tootell, 1996).

In subsequent experiments, we discovered that the lower visual field has better attentional resolution than the upper. The lower visual field is over-represented in the occipital-parietal regions that are often linked to spatial attentional control (Gazzaniga & Ladavas, 1987; Posner et al, 1987). This result again suggests that the locus of the attentional bottleneck lies beyond the primary visual cortex which itself has relatively similar representations of the upper and lower visual fields.

This parietal locus was confirmed by fMRI studies of subjects during attentive tracking tasks (Culham et al, 1997). Moreover, by varying the number of items to be tracked attentively we were able to pinpoint the parietal structures that were involved in attentional control as compared to structures that were activated by the task but indifferent to the degree of attentional load (such as the frontal eye fields which perhaps control steady eye fixation).

AN ALTERNATIVE ROLE FOR ATTENTION

Why would the visual system encode details at resolutions far finer than attention can process? One answer is that attention is not the only client for visual information. There are other equally important paths taken by visual information contributing to motor control for example. I will call the core visual system which services all of these clients the *primary visual system*. A widely accepted view of vision supposes that a stable representation of the 3-dimensional world is built up by adding simple details effortlessly and integrating more complex details with the help of attention. Once constructed, this internal model of the world is considered to persist without additional effort. The following section describes several challenges to this view and suggests an alternative. The ideas presented are closely related to those of Baars (1988) concerning the public broadcast metaphor of consciousness and to Goodale and Milner (1992) concerning the role of alternative visual systems whose content is unavailable to awareness.

I will argue that the selection by attention is not the final stage of vision, but is a separate branch of output for visual information, one that is intended for "public broadcast" (Baars, 1988) and one that has its own constraints which are not necessarily representative of the overall visual analysis of events. Since most experiments on vision are based on a subject's reports (i. e. memory) of their conscious percepts, the large body of information concerning vision has mainly tapped only this separate branch of conscious perception, the experience of attended and therefore remembered events.

CONSCIOUS PERCEPTION IS SPARSE

Recent work has dramatically altered how we think of the consciously experienced world. Rather than rich and complete, it seems that it is sparsely represented and constantly shifting. Many agree that attention is the gateway to conscious perception. For example, Mack and Rock (1996) have demonstrated that events which would otherwise be easily seen go unnoticed if they are unexpected, even when presented at the center of gaze. Joseph, Chun, and Nakayama (1996) have shown that events which have been claimed to require no attentional processing -- the 'pop-out' of a tilted line among vertical lines -- go undetected if attention is diverted.

But what of our feeling that we piece together our world in multiple glances, building up a reasonably complete model of a stable world around us? This concept of a rich model of the world does not hold up. Rensink and his colleagues (Rensink, O'Regan, & Clark, 1996) have shown that large dramatic changes to images go unnoticed unless they occur to objects currently under scrutiny by attention. Simons and Levin (1997) have made the same point with scene changes in video and movie editing. As the camera angle changes from one view to another, a change in clothing or an object or even replacing one person with another goes unnoticed by the majority of observers.

These changes would not be missed if we did have a model of our surroundings that was as complete as we feel it to be. It would be disquieting if, in fact, the world changed every time we blinked, if someone new was substituted for our waiter every time he or she appeared. But these studies have demonstrated that we would not notice these changes even if they were occurring regularly. Luckily, the world is not so unpredictable but this stability is a property of the world, not our representation of it.

Not only is attention necessary to select the elements for this sparse, conscious description of the scene but it appears that it is required to maintain the descriptions as well (Wolfe, Klempen, & Dahlen, 1998). If attention wanders off elsewhere, the items so laboriously constructed are no longer available to conscious inspection.

If rapid, mental handiwork is required just to create a *false* sense of "knowing what is out there", we can again ask how is it we so effortlessly and accurately navigate around our world, sitting down on chairs, picking up forks, driving while talking. Some authors have proposed that the motor systems have their own simple visual representation of the world and that this is sufficient to support simple motor behavior (Goodale & Milner, 1992). I suggest that it is conscious perception which is the supplementary, condensed visual system. The degree of representation within the primary visual system itself (as opposed to conscious perception) may be much more extensive, affording the support for motor behavior and the descriptions of objects and events which can be picked up rapidly by conscious perception when the need arises.

EXPORTING VISION TO THE BRAIN

Conscious perception can be viewed as a description of the visual world exported to the rest of the brain -- it is not necessarily a good representation of the information generated by the visual system, it is just one, strongly constrained description of it. If we accept the relatively modular nature of the brain, where the frontal lobes, the motor areas and the sensory regions have a good deal of independence, it is clear that coordinated activity requires some form of communication among these modules. The modules must talk to each other and it is reasonable to assume that each module will make announcements of its events available to all other modules rather than sending specially tailored notes to each (e. g. Fodor, 1983). This broadcast function or "public billboard" (Baars, 1988) requires a common language for exchanges, a common format interpretable by all the participating modules.

Taking the case of the visual system as an example, I propose that a major function of attention is the translation of representations generated in the visual system into this public language. Some events in the visual representation are structured so as to be very easily and rapidly translated into the public format and posted on the billboard of the mind. Other events require substantial processing to be translated. The assumption here is that the visual representation in primary vision is highly elaborated without the intervention of attention. Attention serves only to read out requested portions of the representation and a translate them into public format.

In other words, the difficulty of locating conjunctions in visual search, for example, does not mean that attention is required to integrate features. An integrated description of many visual properties may be generated in the primary visual system without the intervention of attention. It is just that many of these descriptions may not easily be read out in the available "terms" of the public code and therefore have to be laboriously translated into the best approximation under attentional control.

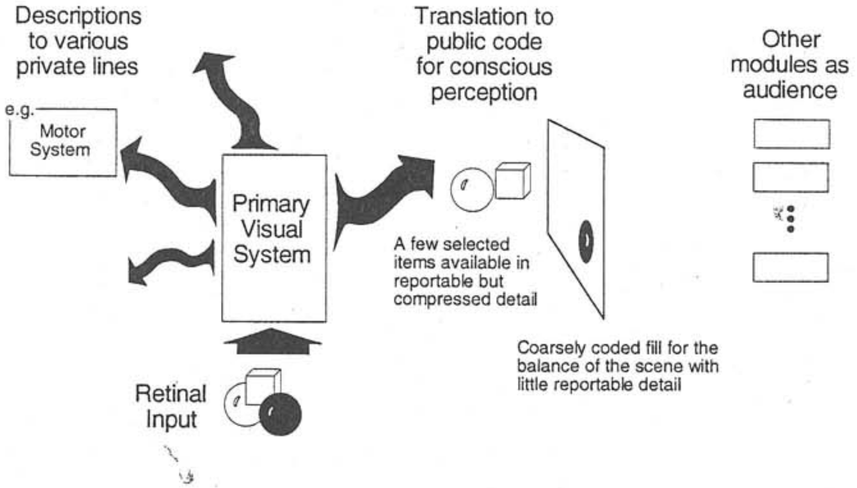

Figure 3. The primary visual system generates elaborated descriptions of the visual world and exports these to various " clients" . The description exported to conscious perception must be coded in a public format to be understandable to all the modules receiving the public broadcast. The role of attention is to select descriptions, translate them and maintain them on the public billboard. Many of the properties of vision may be properties of this common code; the representations of the primary visual system may be less constrained. This sketch has only noted the flow of information from the retina to other centers. Undoubtedly, information is flowing in the other direction as well.

ATTENTION-BASED VISUAL ROUTINES: SPRITES

It is possible that some levels of description computed under attentional control are not available in any form to the primary visual system. In this case, no automated "private line" connection could ever emerge for these aspects of visual representation, attention and awareness would always be required.

Simple examples are counting of elements in a display or deciding whether a dot is inside or outside a closed contour (Ullman, 1984). The processes we have studied in most detail in our lab include attentive tracking (Pylyshyn & Storm, 1988) of random and constrained motion displays (like point-light walkers). We have proposed that attention tracking supports an entire motion analysis on its own (Cavanagh, 1992) and studied the parietal activation that accompanies tracking using fMRI techniques (Culham et al, 1997).

CONCLUSIONS

This paper proposed that attention has very coarse spatial access and that one of its central roles is to translate the description available in the primary visual system into the formats used for public broadcast to the rest of the mind. Much of the measurement of the deployment of attention may reflect properties of this translation step rather than properties of the underlying visual representation.

ACKNOWLEDGMENTS

This research was supported by NIH grants to PC and SH.

REFERENCES

Baars, B.J. (1988). *A cognitive theory of consciousness*, Cambridge, UK: Cambridge University Press.

Campbell, F.W., and Gubisch, R.W. (1966). Optical quality of the human eye. *Journal of Physiology, 186*, 558-578.

Castiello, U., and Umilta, C. (1992) Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 837-848.

Cavanagh, P. (1992). Attention-based motion perception. *Science, 257*, 1563-1565.

Crick, F., and Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature, 375*, 121-123.

Culham, J. C., Cavanagh, P., Kanwisher, N. G., Brandt, S., Dale, A. M., and Tootell, R. B. H. (1997). Attentive tracking of moving targets produces parietal activation revealed by functional magnetic resonance imaging. *Investigative Ophthalmology & Visual Science, 38*, S1174.

DeYoe, E. A., & van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in the Neurosciences, 11* 219-226.

Eriksen, C.W.W., and Hoffman, J.E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics, 12*, 201-204.

Fodor, J.A. (1983). *The modularity of the mind: An essay on faculty psychology*, Cambridge, MA: MIT Press.

Gazzaniga, M.S., and Ladavas, E. (1987). Disturbances in spatial attention following lesion or disconnection of the right parietal lobe. In Jeannerod, M. (Ed.), *Neurophysiological and neuropsychological aspects of spatial neglect, Vol. 45* (pp. 203-213). Amsterdam, Netherlands: Elsevier Science Publishing Co, Inc.

Goodale, M.A. and Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences , 15*, 20-25.

He, S., Cavanagh, P., and Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*, 334-337.

Joseph, J.S., Chun, M.M., and Nakayama, K. (1996). Attention requirements in a 'preattentive' feature search task. *Nature, 387*, 805-807.

Koch, C., and Tootell, R.B. (1996). Visual psychophysics: stimulating brain but not mind [news; comment]. *Nature, 383*, 301, 303.

Kramer, A.F., and Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science, 6*, 381-386.

Mack, A., and Rock, I. (1996). Inattentional blindness: Perception without attention. In Wright,

R. D. (Ed.), *Visual Attention.* New York: Oxford University Press.

Movshon, J.A., and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature, 278*, 850-852.

Posner, M.I., Walker, J.A., Friedrich, F.A., and Rafal, R.D. (1987). How do the parietal lobes direct covert attention? Special Issue: Selective visual attention. *Neuropsychologia, 25*, 135-145.

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

Rensink, R., O'Regan, J.K., and Clark, J.J. (1996), To see or not to see: The need for attention to perceive changes in scenes. *Investigative Ophthalmology & Visual Science, 37*, S213.

Simons, D.J., and Levin, D.T. (1997). Change blindness. *Trends in Cognitive Science*, in press.

Ullman, S. (1984). Visual routines. *Cognition, 18*, 97-159.

Ungerleider, L. G. & Mishkin, M. (1982). Two cortical systems. In D. J. Ingle, M. A. Goodale and R. J. W. Mansfield (eds.) *Analysis of Visual Behavior*, Boston: MIT Press, 549-586.

Wolfe, J.M., Klempen, N., and Dahlen, K. (1998). Post-attentive vision. submitted.