INSIGHTS ABOUT SEEING

Review of *Sight Unseen: An Exploration of Conscious and Unconscious Vision*, by Melvyn A. Goodale and A. David Milner. ISBN 0-19-856807-X, Oxford: Oxford University Press, 2005, 146 pages. Price (softcover) US $29.50, UK £ 14.99.

Goodale and Milner's new book *Sight Unseen* is a beautifully written outline of their view of human vision. The descriptions of how the famous patient DF deals with everyday problems give a sense of her visual abilities and disabilities that cannot be obtained from the many scientific papers about her. But the book is more than a scientific biography of DF. It is a plea for the hypothesis that the human brain has two quite separate visual systems: one for handling our perception and another for guiding our actions. Arguments based on studies with DF and several other neurological patients are supported with arguments based on behavioural research using subjects with normal vision and on arguments based on knowledge of the anatomy and physiology of the visual system. The book is fun to read and very convincing. In fact, the authors' ideas have proven to be so convincing that the two visual systems hypothesis already has a tremendous influence on how scientists think about vision. But are the arguments for two separate streams of visual information really justified?

In order to judge this we have to consider exactly what the claim is. In the first two chapters Goodale and Milner describe the selectivity of DFs visual problems. They very convincingly illustrate the contrast between her very limited visual experience of the world and her almost normal visual control of her own actions. Together with the description of patients with complementary deficits in chapter 3, it becomes completely clear that the human brain must be divided into spatially and functionally segregated pathways. And this is not only so for the human brain. In chapter 4 work is described that shows that frogs have at least two visuo-motor pathways: one for catching prey and another for avoiding obstacles. These two pathways have been studied in detail, but there is no reason to believe that this segregation is unique. Goodale and Milner propose that vision has evolved "as an expanding collection of relatively independent visuo-motor modules" (page 44). In chapters 5 and 6 evidence and reasons are presented for having many independent modules. In chapter 7 some of the complications and limitations of specialised modules are discussed. All the reasoning is sound and convincing. So how can we not agree with their view? Most of the convincing arguments in *Sight Unseen* are directed against the view that the visual system works towards a single consistent representation of the outside world. Despite its intuitive appeal, this view clearly cannot be upheld. We agree with Goodale and Milner about this idea, as probably do most scientists in this field. There is ample evidence that the brain is built up of many, partially independent modules. But the arguments are less convincing when Goodale and Milner insist on grouping these modules into two pathways, one to generate our percepts and the other for guiding our actions. They even seem to suggest that the pathway that generates our percepts culminates in a single "rich and detailed representation of the visual scene" (page 45). We see no reason for proposing such grouping.

Representations in the Brain

A question that puzzles many people when first learning about how the eye works is how come the world does not look upside-down to us, although the image on the retina is upside-down. Typically, rather than immediately realising that the orientation of the retinal image with respect to gravity is irrelevant, people initially conclude that the brain must turn the image back again. This tendency demonstrates the strength of our intuition that the purpose of vision is to produce a reliable representation of the world. Goodale and Milner appeal to this intuition when they write, "the job of perception, after all, is to construct a useful internal model or representation of the real world outside" (page 82). And this intuition is supported by the fact that the part of our brains that deals with vision is not a mess of randomly connected cells, but is organised into distinct brain areas, within each of which neighbouring cells respond to stimulation of neighbouring (or overlapping) parts of the retina. This reproduction of the spatial layout of the retinal image is even used to define brain areas. So what could be wrong with the claim that the purpose of vision is to produce a reliable representation of the world?

In order to answer this question we need to first look at what we really mean by a representation. Obviously, thinking about a representation as a picture in the brain will not get us anywhere. Who will look at the picture? A representation must contain the processing that is needed to make decisions on the basis of the information. There are many possible decisions, such as what colour or objects one sees, who a person is, where the milk

is, how high a step is, whether ones fingers are moving to the correct points on the cup, and so on. So how many representations do we need for all this? And how are they organised? Apparently frogs have at least 5 different visuo-motor modules (page 42 of *Sight Unseen*). People probably have very many more. As mentioned above, Goodale and Milner appear to be proposing that there is a single *rich and detailed* visual representation somewhere in the perceptual (ventral) stream. This suggestion is unlikely considering our tolerance towards inconsistencies when making visual judgements about related attributes (e.g., Brenner and van Damme, 1999). However this is not Goodale and Milner's main point.

Goodale and Milner's main point is that we need two distinct kinds of representation, because we have two very different kinds of decisions to make. One representation is based on the enduring properties of the scene, and is used for recognising things, people and places. The second representation is based on the instantaneous relationships between objects and us, and is used to guide our actions. Goodale and Milner argue that information is processed differently and separately for these different kinds of tasks, because the tasks need different sorts of information. But does this mean that perception and action need separate pathways?

The Alternative

There is no doubt that visual pathways do segregate. The cells in the many visual areas of the brain differ from each other in terms of the kind of stimuli to which they respond. The segregation probably starts even earlier than Goodale and Milner acknowledge: from the very first synapse in the retina (Calkins and Sterling, 1999). Segregating visual processing as early as possible has the advantage that each pathway can evolve to optimally analyse a certain property. If so, the many different visual areas should not be seen as a single hierarchically organised pathway, or two such pathways, but as many parallel processing streams, with interconnections at the stages at which they need to interact. Interactions can be necessary for all kinds of reasons. Obviously, colour, shape, texture and position may have to be combined if one wants to localise the green apple amongst red apples and green pears. A subtler example is that independently processed retinal disparities and retinal extents may have to be combined with a judgement of distance in order to obtain judgments of an object's shape and size (Brenner and van Damme, 1999). Probably the retinal location is maintained in all the streams because this is the basis for combining the various aspects when necessary.

Thus, a more realistic alternative than working towards a single consistent representation is that the visual system is a collection of miscellaneous lines of processing, each leading to the analysis of a certain aspect of the information that reaches our eyes. The appropriate aspects are then selected and combined for the task at hand. Many of the findings that are presented as evidence for two visual systems are completely consistent with the presence of many channels of processing in the brain. Some arguments that clearly support this alternative view are presented in chapter 5 of *Sight Unseen*, most explicitly on pages 58 and 59. So how does the two visual systems hypothesis differ from this alternative?

Why the Dualism?

The fact that different aspects are processed separately does not exclude the possibility that all these separate aspects are combined somewhere into a single, more abstract representation. In *Sight Unseen* a sound case is presented against this proposal. However, Goodale and Milner take only a very small step away from it. They propose that the separate aspects are combined into two streams. They distinguish between the two on the basis of the goal: perception or action. Of course, if their hypothesis is correct the same distinction can be made in several ways, because of the differences between the kinds of information that you need for recognising the green apple and the kind of information that you need for picking it up. For recognition you will be interested in its colour, the texture on its surface, and its shape. For picking it up you will want to know its position and orientation relative to yourself. The former are enduring properties that are presumably evaluated in the light of your previous experience with such apples. They are the kind of properties that appear to be processed in the temporal lobe. The latter are ever changing properties that are only relevant at that particular moment (Rossetti, 1998). They are more likely to be found in the parietal lobe. So, do these differences mean that there are two distinct pathways, one of which can only guide our actions while the other can only determine what we see, or is the distinction simply that recognition usually relies on other properties than action?

Before turning to the question of what properties are important for different tasks, let us examine the evidence for the proposal that there are only two visual streams. In analogy to the reasoning that led Goodale and Milner to propose that there are separate pathways for perception and action, we can for instance ask whether there are patients who fail to perceive certain properties, but not others. There are quite a few fascinating reports of very specific failures to perceive certain properties: people who can recognise objects but not people, can see objects irrespective of their motion but cannot see the motion, and so on (see page 59 in *Sight Unseen*).

Similarly, in analogy with evidence of illusions influencing perception but not action, one can arrange the Müller-Lyer figure so that the central lines are perceived to have different sizes although the endpoints are perceived to be aligned (see figure 1 in Smeets et al., 2002). Or one can look at a static object after adapting to motion in one direction, in which case the object will appear to move considerably in the opposite direction, but hardly to change its position. Likewise, it is possible for one aspect of an action to be influenced by an illusion whereas another is not (e.g. the acceleration and direction of a hitting movement; Smeets and Brenner, 1995). Thus the same reasoning that is used to argue for two streams of visual processing can be used to show that there are many more streams.

Does Recognition Rely on other Properties than Action

So does the distinction lie in the fact that perception (recognition) relies on other properties than action? This is a much more difficult question to answer than it may seem. The main difficulty arises when one tries to identify the critical distinction between perception and action. As an example, consider an experiment that we conducted several years ago to examine whether action is insensitive to illusions. We asked subjects to lift disks embedded in a Ponzo illusion, and found a clear influence of the illusion on the force that was exerted in order to lift the disks (Brenner and Smeets, 1996). When discussing a replication of this study (page 105), Goodale and Milner (2005) do not hesitate to propose that the grip force is determined in the ventral (perceptual) stream. But does this mean that lifting an object is perception?

The reason that the lift force is influenced by the illusion is clear: the weight is not directly available visually, but must be guessed on the basis of the (mis)perceived size. The question is how this fits in with the proposed dissociation between visual information for perception and action. In our opinion lifting a disk is an action, because we intuitively classify it as such, so this finding is inconsistent with the two visual systems hypothesis. But conceding that lifting is a percept makes matters even worse, because it fundamentally undermines the whole hypothesis. If the distinction between perception and action only becomes evident when one knows what information is used, or whether illusions influence the task, then there is no way to test such a hypothesis. Of course, if we define perception and action on the basis of the information that is used, we will find that they use different information.

In chapter 7, Goodale and Milner propose that the distinction between perception and action is related to consciousness. We often notice a distinction between the use of visual information for recognition and for guiding our actions when we respond to visual information without being aware of it, such as when we avoid colliding with other people while walking down a busy street. Defining perception on the basis of what we notice is consistent with there being only two streams, because there are only two possibilities. So is what we notice dealt with differently than what we do not? Obviously it is dealt with differently at some stage, because something makes us aware of it. But that does not prove that it relies on fundamentally different properties than our unconscious actions. In fact, determining what properties influence our actions is extremely difficult, and may be impossible if we assume that the properties involved are ones that we cannot be aware of. Thus until we either really understand the mechanisms of consciousness, or know how to identify the roots of our actions on the basis of neuronal activity alone, there is no point in discriminating between perception and action in this way either. In fact, considering the many ways in which we can indicate what we perceive, which invariably involve some kind of action (e.g., speaking, pressing a key, moving a computer mouse), and the many complicated ways in which vision can guide what we intuitively classify as actions (e.g., moving a computer mouse to shift a cursor on a screen, walking towards a target), there may not even be a clear distinction.

Returning to DF

Thus we neither see evidence that there are only two visual streams, nor do we see evidence that perception and action rely on fundamentally different information. So how about DF? On reading the beautiful anecdotes in the book we get the impression that DF is unable to combine visual properties (determined in separate areas) into objects. She can clearly tell where things are (so she can grasp and avoid them), and seems to also have access to many of the individual visible properties. But the properties seem not to be combined in the normal manner. It is evident that combining properties is much more critical for recognising objects than for grasping them. That is perhaps why DF is unable to visually recognise objects that she is quite able to grasp. Of course, it is beyond our competence to interpret DF's vision. However, we mention this interpretation in order to illustrate that a different account could explain the findings if we agree to abandon the notion of a single consistent internal representation of space.

Conclusion

In our opinion there is little evidence for a fundamental distinction between visual processing

for perception and for action. However, this should not discourage people from reading S*ight Unseen*. It is great fun to read, and the convincing evidence against a single consistent visual representation in the brain will help non-experts to understand the visual system, by helping them to abandon such intuitions. The two visual systems hypothesis has played an important role in this process, and will probably continue to do so for some time. The book is clearly primarily intended for people not working in this field, for whom the distinctions that we have been trying to make are probably irrelevant. However, we want to warn people not to be too eager to accept the fundamental dualism of the two visual systems hypothesis, but to consider the possibility that the brain is a complex network of very specialised pathways that only interact when they need to.

Eli Brenner and Jeroen B.J. Smeets

References

Brenner E and Smeets JBJ. Size illusion influences how we lift but not how we grasp an object. *Experimental Brain Research, 111:* 473-476, 1996.

Brenner E and van Damme WJM. Perceived distance, shape and size. *Vision Research, 39:* 975-986, 1999.

Calkins DJ and Sterling P. Evidence that circuits for spatial and color vision segregate at the first retinal synapse. *Neuron, 24:* 313-321, 1999.

Rossetti Y. Implicit short-lived motor representations of space in brain damaged and healthy subjects. *Consciousness and Cognition, 7:* 520-558, 1998.

Smeets JBJ and Brenner E. Perception and action are based on the same visual information: Distinction between position and velocity. *Journal of Experimental Psychology: Human Perception and Performance, 21:* 19-31, 1995.

Smeets JBJ, Brenner E, de Grave DDJ and Cuijpers RH. Illusions in action: Consequences of inconsistent processing of spatial attributes. *Experimental Brain Research, 147:* 135-144, 2002.

Eli Brenner and Jeroen B.J. Smeets, Faculty of Human Movement Sciences, Vrije Universiteit Van der Boechorststraat 9 NL-1081 BT Amsterdam, The Netherlands.
e-mail: e.brenner@fbw.vu.nl
http://www.fbw.vu.nl/persona/Brenner.htm
e-mail: j.smeetsr@fbw.vu.nl
http://www.fbw.vu.nl/~JSmeets/