ACKNOWLEDGMENTS
BJS was supported by NIMH grant No. R03-MH63808-01. DJS was supported by NSF grant No. BCS-9905578 and by an Alfred P. Sloan Research Fellowship.

The absence of representations causes inconsistencies in visual perception

Jeroen B. J. Smeets and Eli Brenner
Vakgroep Fysiologie, Erasmus Universiteit Rotterdam, NL-3000 DR Rotterdam, The Netherlands. **smeets@fys.fgg.eur.nl**
brenner@fys.fgg.eur.nl **www.eur.nl/fgg/fys/people/smeets.htm**

Abstract: In their target article, O'Regan & Noë (O&N) give convincing arguments for there being no elaborate internal representation of the outside world. We show two more categories of empirical results that can easily be understood within the view that the world serves as an outside memory that is probed only when specific information is needed.

In line with the arguments in the target article, we consider vision to be tightly coupled to motor control. In order to catch a ball, one needs information about its size, weight, position, speed, and direction of motion. These attributes are important for different aspects of the action, so that they can be determined and processed independently within what has become known as separate visuomotor-channels (e.g., Jeannerod 1999).

Although determining visual attributes independently might be useful for controlling actions, this does not mean that the outcomes are independent, because the laws of physics and geometry relate many of these attributes. For instance, if an object moves at a certain speed, its position will change at a corresponding rate. An internal representation of the outside world would combine all available information to yield the most likely (and thus consistent) representation. This would of course reflect the physical and geometrical relationships within the outside world. The consequence of independent processing is that the relevant sources of information are combined separately for each attribute. Physically related attributes might thus be determined on the basis of different sources, within physiologically independent pathways. If all attributes are determined veridically, this independence remains unnoticed. It becomes evident when the processing of one attribute is erroneous, as is the case in visual illusions (Smeets & Brenner 2001). Two examples clarify this.

For intercepting a moving object one needs information about its speed to regulate the timing of one's action, and information about its (egocentric) position to direct one's action. Due to the noisiness of extraretinal information on eye orientation, the most accurate estimate of object speed will generally be one based on relative retinal information (Smeets & Brenner 1994). For determining an object's egocentric position, the use of extraretinal information cannot be avoided. And indeed, moving a visual background influences the perceived speed, without influencing the perceived position (Duncker illusion). In our view, each such attribute is processed independently to control a certain aspect of our actions. The Duncker illusion therefore affects the timing of ones action, without influencing its direction (Smeets & Brenner 1995).

A similar reasoning holds for grasping an object to pick it up. To move the digits to the object's surface, information about positions on that surface is needed (Smeets & Brenner 1999). To subsequently apply adequate forces to lift the object, a visual correlate of the object's weight is needed: that is, its size (Gordon et al. 1991). As with the previous example, these geometrically related aspects (positions and size) might very well be determined on the basis of different sources of information. The positions will again be determined using extraretinal information, whereas the object's size might be determined purely on the basis of retinal information. This explains why illusions of size affect the lifting force in grasping, but not the grip aperture (Brenner & Smeets 1996).

Independent processing of physically related attributes is not only evident in the visual control of action, but also in conscious perception. For instance, if one looks for a while at a waterfall, and subsequently fixates a tree at eye-level near that waterfall, the tree appears to move upward. The apparent position of the tree remains approximately at eye-level. Other examples of inconsistencies can be found in visual illusions, such as the Müller-Lyer illusion. This illusion influences the perceived size of the figure without affecting the perceived positions of the end-positions (Gillam & Chambers 1985). In analogy to the claim that we process only one fragment of the world at a time (sect. 4.2), this apparent inconsistency suggests that conscious perception involves processing only one attribute of that fragment at a time.

If one accepts that not all attributes are processed at a time, one can understand the flash-lag effect (e.g., Nijhawan 1994). This effect manifests itself when a subject is fixating a screen on which a target is moving continuously while another target flashes. If the subject is asked to indicate the position of the moving target at the time of the flash, he will misjudge this position in the direction of the target's motion. This has been interpreted as the result of motion extrapolation. However, this cannot be so because if the target unexpectedly reverses direction near the moment of the flash, the misjudgements are never beyond the actual trajectory of the moving target. It is more likely to be caused by different processing times for flashed and continuously presented stimuli (Whitney & Murakami 1998).

However, there is no reason to assume that flashes are processed more slowly than continuously visible stimuli. What then is the cause of this apparent difference in processing time? If not all attributes are processed continuously, the position of the moving target will have to be probed at some instant. This presumably takes time, and can start only after the flash has been detected. The moving target's position (or other attributes such as its colour and shape) will be probed too late. If this explanation is correct, the flash-lag effect should disappear if we change the experiment in a way that allows the position of the moving target to be probed at the time of the flash. A simple way to do so is to provide an additional cue for the time (or equivalently, the position of the moving target) at which the flash will occur. Indeed, the flash-lag effect is reduced markedly when this is done (Brenner & Smeets 2000).

Re-presenting the case for representation

Benjamin W. Tatler
School of Biological Sciences, University of Sussex, Brighton, BN1 9QG United Kingdom. **b.w.tatler@sussex.ac.uk**
http://www.biols.susx.ac.uk/resgroups/scn/vision

Abstract: O'Regan & Noë (O&N) present the most radical departure yet from traditional approaches to visual perception. However, internal representation cannot yet be abandoned. I will discuss: (1) recent evidence for very short-term pictorial representation of each fixation; (2) the possibility of abstract representation, largely unconsidered by the authors; and (3) that sensorimotor contingency theory requires internal visual retention and comparison.

O'Regan & Noë (**O&N**) extend the implications of recent change detection studies by arguing that not only is it unnecessary for the visual system to construct a point-by-point pictorial representation of the world across multiple fixations, but that no such information need be internalised on even the shortest of time scales. However, the reader should be cautious before abandoning all notion of representation and should first consider some of the implications of this model and other possible accounts.

Whilst it would be hard to argue that we build up a point-by-