Visual Routes to Action and Perception

Nicholas Cothros

The present review considers how visual input is used by the nervous system for perception or to guide action. While all visual input arises from the eyes, the information carried by neural pathways from the eyes to the brain diverges. This divergence can be thought of in terms of a bifurcation, with some visual information being used in the service of perception (as in the recognition of objects, faces, scenes and so on) and the rest being used to direct the motor system (as in the appropriate scaling of grip aperture for grasping, or the selection of a target for pointing or eye movements and so on).

The eye. Before considering how the nervous system interprets visual input, it is necessary to consider how light is transduced into neural input. The organ responsible for light detection is the eye. It is essentially a chamber with a hole and lens at one end, for receiving and focusing light, and a screen at the other end for the projection of light.

The cornea is part of the outermost layer of the eye. Situated just outside the pupil (and forming the anterior chamber), it is convexed and responsible for most of the refraction of light entering the eye. Light enters through the pupil, which is simply a hole in the eye. The iris is a drawstring-like muscle which can increase or decrease pupil size and control the amount of light entering the eye. Beyond the pupil lies a lens, actuated by small ciliary muscles. The lens further bends incoming light, converging or diverging rays, depending on its shape (partly determined by the aforementioned small muscles).

Behind the lens lies the posterior chamber of the eye. It is a space filled with vitreous humour, which is more dense and gelatinous the the aqueous humour of the anterior chamber. The posterior end of this chamber is lined with a layer of cells. This area, called the retina, is the projection screen of the eye.

The retina. Light entering the eye strikes the retina, which comprises a number of different cell types. The photoreceptors are the first link in a chain of interconnected cells. Photoreceptors are sensitive to light and are broadly of two types: the rods and the cones. Cones are primarily for colour detection and function well in bright light. Rods are more sensitive to light, but do not detect colour well and so they are adapted for vision in dim light. In humans, the retina contains roughly 120 million rods and 6 million cones.

Light striking a photoreceptor results in a chemical reaction (mediated by photopigments, like rhodopsin in the rods), which in turn leads to a change in the cell's membrane potential. Photoreceptors do not generate action potentials. They do generate graded potentials. There are three types of cones (each with a different photopigment), which differ in terms of the light wavelengths they are sensitive to. This forms the basis of colour vision.

Photoreceptors hyperpolarize to light. Light effectively shuts them "off". It is primarily through the action of networks of inhibitory horizontal cells that signals from photoreceptors are modified. This input ultimately arrives at retinal ganglion cells, which produce action potentials. As indicated earlier, neural input from photoreceptors is modified through horizontal cells. The end result of this is that ganglion cells can respond either to light onset or light offset. This relates to the concept of the "receptive field".

The receptive field. In a sensory neuron (like the retinal ganglion cell), the receptive field is an area in which the presence of a stimulus (in this case, light on the retina) alters the firing of said neuron. The receptive fields of ganglion cells are of two types, which shall be

referred to as "on-centre" and "off-centre". An on-centre ganglion cell is stimulated when light hits the centre of its receptive field; it is inhibited when the part of the receptive field *surrounding* the centre is exposed to light. Stimulation of the centre of the receptive field increases the firing rate of the ganglion cell. Stimulation of the surround decreases the firing rate of the cell. Stimulation of both the centre and surround produces a response of intermediate strength. Off-centre ganglion cells work in the opposite way.

The on-centre or off-centre arrangement of ganglion cells' receptive fields is the product of lateral inhibition via inhibitory horizontal cells. Ganglion cells transmit information concerning the differences in firing rates of cells in the centre versus those in the surround of the receptive field. Ganglion cells thus transmit information about contrast.

Receptive fields vary in size. Ganglion cells with small receptive fields are optimally stimulated by high spatial frequencies (in other words, visual stimuli with fine detail). Ganglion cells with large receptive fields are optimally stimulated by low spatial frequencies, which pertains to visual stimuli with coarse detail. One can imagine how the edges of objects create interruptions in the distribution of light on the retina. These would affect ganglion cell receptive fields (given their on-centre or off-centre arrangement). Thus, ganglion cell activity also pertains to the edges of objects.

Projections to the brain. The output of the ganglion cells exits the retina through the optic nerve and synapses on the lateral geniculate nucleus (LGN) of the thalamus. A division of visual input begins here, not only because layers of the LGN independently represent each eye, but because the LGN layers are also separated into magnocellular and parvocellular layers. These two respond to rather different types of visual input. The magno layers are innervated by large ganglion cells with large receptive fields. These magno cells respond to fast, transient visual stimuli and are most sensitive at low contrast. The parvo layers are innervated by small ganglion cells with small receptive fields. These parvo cells respond to slow, sustained stimuli and are most sensitive at high contrast.

From the LGN, visual input proceeds to the primary visual cortex, also known as striate cortex or V1. The receptive fields of cells in V1 are tuned for more complex stimuli than those of cells in the retina. The receptive fields of so-called "simple" cells in V1 are optimally stimulated by bars of light in a specific orientation. This is noted in the columnar organization of V1, which is divided into orientation columns (along with spatial frequency and ocular dominance columns).

The two visual streams. From V1, the division of visual input continues. Generally speaking, visual input flows from V1 to two distinct areas in the cerebral cortex. These are referred to as "visual streams". In particular, there is a dorsal stream, which corresponds to visual input moving from V1 to dorsal parietal areas, and there is a ventral stream, which corresponds to visual input moving from V1 to inferior temporal areas. Evidence suggests that the ventral stream is involved in using visual input for perception, while the dorsal stream uses vision for action. Thus, the brain appears to have two distinct visual systems.

This is evident when one considers the results of lesions to the dorsal and ventral streams. Damage to the dorsal stream can result in optic ataxia. In this condition, movements that require vision, as in the visual identification and selection of a target, are impaired. Those suffering from this condition grasp for objects as though blind, though their vision appears otherwise normal.

Damage to the ventral stream produces a very different deficit. In this case, the resulting condition is visual agnosia. Patients suffering from visual agnosia cannot identify what they look at. In profound cases, like apperceptive agnosia, patients can neither make sense of

what they're looking at, nor can they copy visual stimuli in drawings. In the less profound, although seemingly more paradoxical case of associative agnosia, patients are able to draw faithful copies of visual stimuli, but perception remains deeply impaired, as these patients cannot identify the visual stimuli. One of the most telling case studies of visual agnosia came from patient DF.

Patient DF. Following carbon monoxide poisoning, DF was left with profound visual form agnosia. While she could roughly draw recognizable items from memory, she was unable to copy visual stimuli resembling those items. She was also unable to recognize such stimuli. It was not DF's deficit, but rather, spared ability in another domain that lent support to the theory that there are two visual systems. DF's behaviour suggests that while her vision for perception is severely impaired, her vision for action is not. While she cannot describe the orientation or shape of objects using verbal report or other descriptive actions, her purposeful motor actions reveal an implicit understanding of the physical features of objects. Naturally, such an understanding is required to guide action, such as grasping. For example, when presented with blocks of varying size, DF's estimates of their size using manual estimation (i.e. splaying the thumb and index finger to pantomime the perceived size of a block) are at chance. However, during a reach towards a block, with the intent to grasp it, DF scales the size of her grip aperture accordingly.

Patient RV. The case of patient RV provides important additional evidence of their being two visual systems. Patient RV suffered damage to the dorsal stream. The behaviour of RV is essentially the reverse of that of DF. In the block task described above, patient RV was able to correctly estimate the size of the blocks using manual estimation. Indeed, RV's ability to recognize and copy visual stimuli matched that of healthy controls. However, it was in using vision for action where RV's deficit became obvious. RV's grasping actions were not scaled according to object orientation. In another task, where subjects were required to establish a stable grip on objects of varying shape, RV's selection of grasp was not precisely matched to the object's shape, despite RV's ability to clearly describe the shape of the objects.

Convergent evidence. Taken together, patients DF and RV provide a double dissociation. The neural substrates of two different abilities have been identified as independent entities. It is possible to lose vision for action but have vision for perception remain intact. It is also possible to lose vision for perception but leave vision for action spared. It is not the case that one type of vision is simply a simpler version of another, as walking is a less strenuous form of leg-based locomotion than running.

Functional neuroimaging has provided additional evidence that there are two visual systems in the brain. Imaging studies have shown that the ventral surface of the brain consists of regions that are particularly sensitive to the presentation of faces (Fusiform Face Area), places (Parahippocampal Place Area) and objects (Lateral Occipital Area). Interestingly, the region of the brain that "lights up" during presentation of objects (area LO) is coincident with the location of DF's lesion.

In the dorsal parietal areas of the brain, functional neuroimaging has revealed areas for grasping (Anterior Intraparietal area), eye movements (Lateral Intraparietal area) and reaching (Parietal Reach Region).

The loci of these action and perception areas in the human dorsal and ventral streams, respectively, also appear to be present in the monkey brain in homologous regions. Electrophysiological recordings in the monkey brain have provided these findings.

Behaviourally, evidence that both visual systems are engaged in healthy controls has been

afforded by the hollow-face illusion. The concave side of a mask resembling a human face, when appropriately lit, elicits a powerful illusion. Specifically, the concave side of the mask appears convex. A recent experiment using the hollow-face illusion entailed generating perceptual estimates of the distance of a point on either the cheek or forehead of the face, or generating purposeful pointing actions towards those same points. The points appeared on either the truly convex side of the face or on the concave side of the face, which was lit to appear convex.

Perceptual estimates of the distance of these points was matched when they appeared on the normal face or the illusory face. These findings indicate that perception of these points was influenced by illusion. However, the same was not observed during the motor portion of this task. When reaching towards these points, the trajectories were scaled to the true distance of the points. In other words, subjects reached further away when aiming towards points on the illusory face. Note that these points truly are at a greater distance, as the illusory face is actually concave and not convex as it appears to be. Also note that visual input was completely blocked by specialized goggles as soon as a reach was initiated, thereby preventing visual feedback from influencing behaviour during action. These results indicate that when generating action, the visual system was not "fooled" by the illusion. In contrast, the formation of visual perception succumbed to illusion.

In addition to providing further evidence of there being two visual systems, the findings of this experiment suggest that actions performed on visible objects do not necessarily rely on the same visual representation that forms the basis of visual perception.

Major Themes/Questions

- What is a receptive field and how does it contribute to visual perception?
- What is the assumption of experience-based control? Is there any evidence suggesting it is false?
- Following a lesion in a particular locus in the visual pathway, from the optic nerve to higher-order visual areas in the cerebral cortex, what deficits can be expected?
- How can one design a neuroimaging experiment to explore the two visual systems? For example, how can one design a neuroimaging experiment to identify the cortical regions involved in reaching versus those involved in grasping?
- What advantages are conferred upon organisms possessing two visual systems?