Perception, Action, and Consciousness

M A Goodale, University of Western Ontario, London, ON, Canada
A D Milner, Durham University, Durham, UK

Glossary

Basal ganglia – Several large clusters of nerve cells (including the caudate nucleus, putamen, and the globus pallidus) located deep in the brain below the cerebral hemispheres. The basal ganglia play a role in a number of functions including motor control, cognition, and emotion.

Blindsight – Residual visual abilities in individuals with large lesions of primary visual cortex who claim that they are blind in that part of the visual field in which they show evidence of visual sensitivity. For example, patients with unilateral lesions of primary visual cortex can sometimes point accurately toward stimuli in visual field contralateral to their lesion while at the same time denying any conscious experience of those stimuli.

Brainstem – The major route by which the forebrain receives information from, and sends information to, the spinal cord and peripheral nerves. The chief divisions of the brainstem are the midbrain, pons, and medulla.

Cortical blindness – Blindness that results from damage to the primary visual areas in the cerebral hemispheres. Individuals who have cortical blindness may nevertheless show blindsight.

Dorsal stream – Visual pathway arising in early visual areas in the cerebral cortex and projecting to the posterior parietal cortex. This pathway is thought to subserve the visual control of actions, such as reaching and grasping.

Extinction – In neuropsychology and neurology, extinction (or bilateral simultaneous extinction) refers to a phenomenon in which a patient with a unilateral lesion (typically in the inferior regions of the posterior parietal cortex) fails to detect a visual (or tactile) stimulus presented contralateral to the lesion when another (similar) stimulus is presented simultaneously on the opposite side. When the contralateral stimulus is presented on its own, the patient is able to detect it.

Hypoxia – A deficiency of oxygen reaching the tissues of the body.

Inferotemporal cortex – The most ventral part of the temporal lobe. The inferotemporal cortex contains higher-order areas of the ventral stream of visual processing.

Metacognition – Awareness of one's own thinking and decision making; sometimes termed 'knowing about knowing.'

Optic ataxia – An inability to guide the hand toward an object using vision. Optic ataxia is caused by damage to the posterior parietal cortex.

Phototaxis – Movement of an organism toward (or away from) light.

Posterior parietal cortex – Cortex in the parietal lobe behind the postcentral sulcus. A prominent sulcus in the posterior parietal cortex is the intraparietal sulcus (IPS). This region of the posterior parietal cortex contains visuomotor areas that make up the dorsal stream of visual processing.

Premotor cortex – A complex mosaic of interconnected areas in the frontal lobe immediately anterior to primary motor cortex. Although there is not complete agreement about the function of different parts of premotor cortex, it has been suggested that areas in this region of cortex participate in motor planning and movement selection. Mirror cells are found in this region.

Ventral stream – Visual pathway arising in early visual areas in the cerebral cortex of the

primate brain and projecting to the inferotemporal cortex. This pathway mediates visual perception, allowing the visual recognition of objects and events. Processing in this pathway is necessary, but not sufficient, for visual awareness.

Visual agnosia – An inability to recognize visual stimuli despite spared low-level visual processing. There are several varieties of visual agnosia, all of which involve damage to some part of the ventral stream. In the case of associative agnosia, the patient is unable to recognize an object despite being able to draw a reasonably faithful representation of what he or she sees. In the case of apperceptive agnosia (or visual form agnosia), the deficit is more fundamental and the patient cannot recognize even simple shapes or discriminate between them – and is unable to copy line drawings.

Introduction

Vision is the primary route to our conscious experience of the world beyond our bodies. Although we certainly hear and feel things in our immediate environment, neither of these sensory experiences is any match for the rich and detailed representation of the world provided by our sense of sight. The majesty of a distant mountain or the angry face of an approaching enemy can be appreciated only as visual experiences. But vision does more than provide us with our perception of the world. It also allows us to move around that world and to guide our goal-directed actions. Although it is tempting to think that these different functions of vision are mediated by one and the same visual representation in our brain, it has become increasingly clear over the last two decades that the visual pathways that underlie our perception of the world are quite distinct from those that underlie the control of our actions. Indeed, the distinction between 'vision-for-perception' and 'vision-for-action' has emerged as one of the major organizing principles of the visual brain, particularly with respect to the visual pathways in the cerebral

cortex. But before elaborating this distinction it is important to understand how vision began.

The Evolutionary Origins of Vision

Visual systems first evolved not to enable animals to perceive the world, but to provide distal sensory control of their movements. One clear example of this is phototaxis, a behavior exhibited by many simple organisms whereby they move toward or away from light. Some bacteria, for example, use orange light as a source of energy for metabolic activity, but must avoid ultraviolet light, which can damage their DNA. As a consequence, these bacteria have developed a differential phototactic response, whereby the system measures light intensity at different wavelengths so that they end up moving toward orange light and away from UV light. To explain the bacteria's light-sensitive behavior, it is not necessary to argue that these single-celled organisms 'perceive' the light or even that they have some sort of internal model of the outside world coded in their one or more of their organelles. One simply has to posit the existence of some sort of input–output device within the bacteria that links the intensity of ambient orange and UV light to the pattern of locomotion. As it turns out, exactly the same argument can be made about the visually guided behavior of more complex organisms, such as vertebrates. Indeed, as we shall see later, a broad range of human behavior can also be explained without reference to experiential perception or any 'general-purpose' representation of the outside world.

In vertebrates, different classes of visually guided behavior have evolved as relatively independent neural systems. For example, in present-day amphibians, such as the frog, visually guided prey-catching and visually guided obstacle avoidance are mediated by separate neural pathways arising in the retina and projecting to distinct motor networks in the brain that produce the constituent movements of these two classes of behavior. In fact, evidence from several decades of work in both frog and toad suggests that there are at least five separate visuomotor modules, each responsible for a different kind of visually guided behavior and each having neural pathways from

input to output. The outputs of these different modules certainly have to be coordinated, but in no sense are they guided by a single general-purpose visual representation in the frog's brain. There is evidence as well that the same kind of visuomotor modularity found in amphibians also exists in the mammalian and avian brain.

Nevertheless, even though there is considerable evidence for visuomotor modularity in all classes of vertebrates, the very complexity of the day-to-day living in many mammals, particularly in higher primates, demands much more flexible organization of the circuitry. In monkeys (and thus presumably in humans as well), many of the visuomotor circuits that are shared with simpler vertebrates appear to be modulated by more recently evolved control systems in the cerebral cortex. Having this layer of cortical control over the more ancient subcortical networks makes it possible for primates to have much more flexible visually guided behavior. But even so, the behavior of primates, particularly with their conspecifics, is so complicated and subtle, that direct sensory control of action is often not enough. To handle these complexities, representational systems have emerged in the primate brain (and presumably in other mammals as well), from which internal models of the external world can be constructed. These representational systems allow primates such as ourselves to perceive a world beyond our bodies, to share that experience with other members of our species, and to plan a vast range of different actions with respect to objects and events that we have identified. This constellation of abilities is often identified with consciousness, particularly those aspects of consciousness that have to do with decision making and metacognition. It is important to emphasize that the perceptual machinery that has evolved to do this is not linked directly to specific motor outputs, but instead accesses these outputs via cognitive systems that rely on memory representations, semantics, spatial reasoning, planning, and communication. In other words, there are a series of cognitive 'buffers' between perceiving the world and acting on it, and the relationship between what is on the retina and the behavior of the organism cannot be understood without reference to other mental states, including those typically described as conscious. But once

a particular course of action has been chosen, the actual execution of the constituent movements of that action are typically carried out by dedicated visuomotor modules not dissimilar in principle from those found in frogs and toads.

To summarize: vision in humans and other primates (and perhaps other animals as well) has two distinct but interacting functions: (1) the perception of objects and their relations, which provides a foundation for the organism's cognitive life and its conscious experience of the world, and (2) the control of actions directed at (or with respect to) those objects, in which separate motor outputs are programmed and controlled online. These different demands on vision have shaped the organization of the visual pathways in the primate brain.

Two Visual Systems in the Primate Cerebral Cortex

In the primate brain, two 'streams of visual processing' arise from early visual areas in the cerebral cortex and project to higher-order visual areas (**Figure 1**). One of these projection systems, the
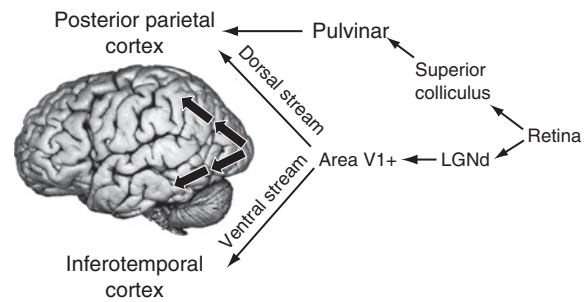
Figure 1 Schematic representation of the two streams of visual processing in human cerebral cortex. The retina sends projections to the dorsal part of the lateral geniculate nucleus in the thalamus (LGNd), which projects in turn to primary visual cortex (V1). Within the cerebral cortex, the ventral stream arises from early visual areas (V1+) and projects to regions in the occipito-temporal cortex. The dorsal stream also arises from early visual areas but projects instead to the posterior parietal cortex. The posterior parietal cortex also receives visual input from the superior colliculus via the pulvinar. On the left, the approximate locations of the pathways are shown on an image of the brain. The routes indicated by the arrows involve a series of complex interconnections.

dorsal stream, projects from early visual areas to the posterior parietal cortex, a region of the brain that is reciprocally connected to motor areas in the frontal cortex and sends projections to the basal ganglia and other (older) motor nuclei in the brainstem. The dorsal stream also receives visual input from the midbrain (via the thalamus). The other projection system, the ventral stream, projects from early visual areas to inferotemporal cortex. The ventral stream has strong connections with medial temporal areas, including the amygdala and hippocampus, as well as the prefrontal cortex. As it turns out, the functions of the two streams map quite well onto the distinction between vision-for-action and vision-for-perception discussed above. Not surprisingly, given its interconnections with motor structures, it is the dorsal stream that plays the critical role in the programming and control of actions, transforming real-time information about the location and disposition of objects into the coordinate frames of the relevant motor systems. In contrast, it is the ventral stream (together with associated cognitive networks) that mediates the construction of the rich and detailed representations of the world that allow us to identify objects, events, and actions in others, attach meaning and significance to them, and infer their causal relations. In summary, processing in the ventral stream provides the conscious visual percepts that are essential for accumulating a knowledge base about the world, knowledge that we can access for cognitive operations, such as planning and decision making. Processing in the dorsal stream does not generate visual percepts; it generates skilled actions (as part of a network of structures involved in sensorimotor control). Of course, the two streams are not hermetically sealed from one another. Indeed, they work together in controlling our behavior as we live our complex lives – but they play separate and complementary roles in the production of adaptive behavior.

p0035   Much of the evidence for this idea first came from work with neurological patients. The best known of these cases is the patient D.F., who developed a profound 'visual form agnosia' as a consequence of a hypoxic episode in which her brain was starved of oxygen. The nature of D.F.'s deficit in form vision can be understood to some extent by examining the drawings illustrated in

Figure 2. Not only was D.F. unable to identify the simple line drawings illustrated in left-hand column of this figure, but she was also unable to copy them, at least in a recognizable way. A preserved ability to see fine detail allowed her to depict some aspects of the drawings, such as the dots indicating the print in the line drawing of the open book. Nevertheless, she was unable to duplicate the overall shape or arrangement of the elements of the line drawings in her copies. D.F.'s inability to copy the drawings is not due to a problem in controlling the movements of the pen or pencil; when she was asked on a separate occasion to draw an object from memory, she was able to do so reasonably well, as the drawings on the right-hand side of **Figure 2** illustrate. Needless to say, when D.F. was shown any of the drawings she had done herself, whether the ones retrieved from memory or those copied from another drawing, she had no idea what they were and commented that they all looked like 'squiggles.' It is important to emphasize that D.F. retains the ability to perceive and describe the colors and other surface properties of objects, such as their visual texture. The missing elements in her perceptual experience are restricted to shape and form, and
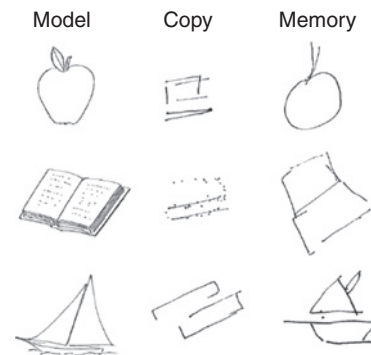
Model   Copy   Memory

Figure 2   The patient D.F.'s attempts to draw from models and from memory. D.F. was unable to identify the line-drawings of the apple, open book, or the boat shown on the left. In addition, her copies were very poor. Note, however, that she did incorporate some elements of the line-drawing (e.g., the dots indicating the text in the book) into her copy. When she was asked on another occasion to draw these same items from memory, she produced a respectable representation of all three (right-hand column). When she was later shown her own drawings, she had no idea what they were.

thus her problems cannot be dismissed as a generalized inability to make perceptual reports.

Remarkably, even though D.F. shows no perceptual awareness of the form or dimensions of objects, she automatically adjusts her hand to the size, shape, and orientation of an object as she reaches out to pick it up. For example, even though D.F. is unable to distinguish between rectangular blocks of different dimensions, when she reaches out to pick up one of the blocks, the aperture between her fingers and thumb is scaled in-flight to the width of the object, just as it is in people with normal vision (see **Figure 3a**). In other words, D.F. can scale her grip to the dimensions of an object in anticipation of picking it up, even though she is unable to 'perceive' the dimensions of that object. Similarly, she will rotate her hand correctly for objects placed in different orientations, and will direct her fingers to stable grasp points on the surface of the objects, even though in other tests she fails to indicate either verbally (or in a manual matching task) the orientation or shape of those same objects. She is also able to avoid other objects in the workspace as she reaches out toward a goal, even though she cannot judge their relative locations correctly in a more perceptual task. D.F. exhibits normal visuomotor control in other tasks

as well, including stepping over obstacles during locomotion, despite the fact that her perceptual judgments about the height of these obstacles are far from normal. To summarize: despite the fact that D.F. has lost all conscious perception of the form of objects, including their size, shape, and orientation, her visuomotor systems are able to make use of these same object properties to control skilled object-directed actions.

So where is the damage in D.F.'s brain? Although D.F. shows some diffuse loss of tissue throughout her cerebral cortex (consistent with hypoxia), she has prominent focal lesions bilaterally in a region of the human ventral stream that has been shown to be involved in the visual recognition of objects. It is presumably the damage to these object-recognition areas that has disrupted her ability to perceive the form of objects. But clearly these ventral-stream lesions have not interfered with her ability to use visual information about form to shape her hand when she reaches out and grasp objects. The preservation of normal visually guided grasping in the face of ventral-stream damage suggests this ability is dependent on another visual pathway, the most likely candidate being the dorsal stream. This conclusion has been supported by neuroimaging evidence showing that when D.F. grasps
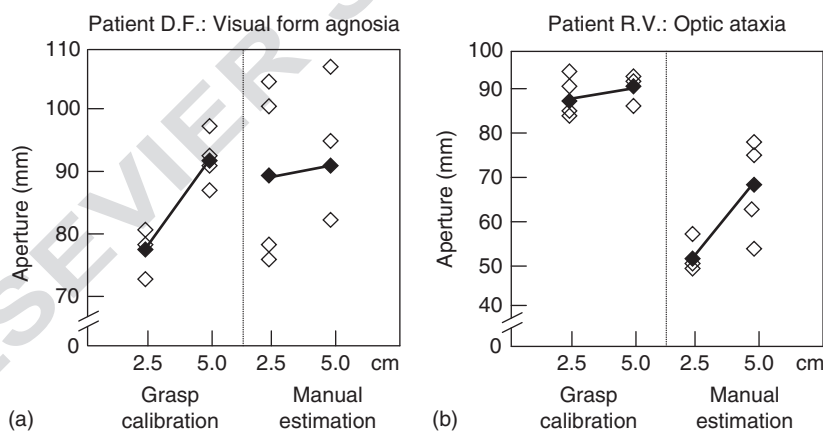
Figure 3 Graphs showing the size of the aperture between the index finger and thumb during object-directed grasping and manual estimates of object width for D.F., a patient with visual form agnosia and R.V., a patient with optic ataxia. D.F. (a) showed excellent grip scaling, opening her hand wider for the 50 mm-wide object than for the 25 mm wide object. D.F.'s manual estimates of the width of the two objects, however, were grossly inaccurate and showed enormous variability from trial to trial. In contrast, R.V. (b) was able to indicate the size of the objects reasonably well (individual trials marked as open diamonds), but her maximum grip aperture in flight was not well-tuned. She simply opened her hand as wide as possible on every trial. Reproduced from Goodale MA, Milner AD, Jakobson LS, and Carey DP (1991) A neurological dissociation between perceiving objects and grasping them. *Nature* 349: 154–156, with permission from Nature Publishing Group.

objects that vary in size and orientation, she shows relatively normal activity in a small region of the dorsal stream that has been implicated in the visual control of grasping in healthy individuals. It should be emphasized that although patients like D.F. are rare, there are a number of other cases in the literature in which the same striking dissociation between visual perception and visuomotor control has been documented.

But what about patients with damage to the dorsal stream? As it turns out, these patients exhibit a pattern of deficits and spared abilities that are complementary to that seen in patients with ventral-stream lesions. Thus, patients with dorsal-stream lesions typically have problems reaching toward targets placed in different positions in the visual field, particularly the periphery. This deficit is referred to clinically as 'optic ataxia.' But the failure to locate an object with the hand cannot be construed as a problem in spatial representation: many optic ataxia patients, for example, can describe the relative position of the object in space quite accurately, even though they cannot direct their hand toward it. Also, the deficit is not purely motor: these patients usually have no difficulty using input from other sensory systems, such as proprioception or audition, to guide their movements. In addition to their deficits in reaching, many patients with damage in the dorsal stream are unable to use visual information to rotate their hand, scale their grip, or configure their fingers properly when reaching out to pick up objects, even though they are able to correctly report the orientation, size, and shape of those objects (see **Figure 3b**). In addition, they do not take into account the positions of potential obstacles when they are attempting to reach out toward goal objects even though they can indicate the relative location of the obstacles in other ways. In summary, patients with optic ataxia exhibit neither a purely visual nor a purely motor deficit, but instead a specific deficit in visuomotor control, confirming the critical role that the dorsal stream plays in the control of skilled actions.

In addition to such work with neurological patients, there is a wealth of evidence from monkey neurophysiology and human neuroimaging supporting the idea of a ventral 'perception' stream and a dorsal 'action' stream.

Different Metrics and Frames of Reference for Perception and Action

But why did two separate streams of visual processing evolve in the primate cerebral cortex in the first place? Or, to put it another way, why couldn't one 'general purpose' visual system handle both vision-for-perception and vision-for-action? The answer to this question lies in the computational requirements of the two kinds of vision. To be able to grasp an object successfully, for example, the brain must compute the actual size of the object, and its orientation and position with respect to the grasping hand of the observer (i.e., in egocentric coordinates). The time at which these computations are performed is equally critical. Observers and goal objects rarely stay in a static relationship with one another and, as a consequence, the egocentric coordinates of a target object can often change radically from moment to moment. For these reasons, it is essential that the required coordinates for action be computed in an egocentric framework at the very moment the movements are to be performed.

The computations underlying perception are quite different. Vision-for-perception does not deliver the absolute size of objects or their egocentric locations. In fact, such computations would be counterproductive for a recognition system precisely because we almost never stay fixed in one place in the world. The problem can be easily solved by the alternative strategy of encoding the size, orientation, and location of objects relative to each other. Such a scene-based frame of reference permits a perceptual representation of objects that transcends particular viewpoints, while preserving information about spatial relationships (as well as relative size and orientation) as the observer moves around. Indeed, it has been suggested that if the perceptual machinery had to deliver the real size and distance of all the objects in the visual array, the computational load would be prohibitive. The products of perception also need to be available over a much longer timescale than the visual information used in the control of action. We may need to recognize objects we have seen minutes, hours, days – or even years before. In short, the perceptual information is lodged in memory. To achieve this, the coding of the visual information has to be

somewhat abstract – transcending particular viewpoint and viewing conditions. By working with perceptual representations that are object- or scene-based, we are able to maintain the constancies of size, shape, color, lightness, and relative location, over time and across different viewing conditions. Although there is much debate about how this is achieved, it is clear that it is the identity of the object and its location within the scene, not its disposition with respect to the observer that is of primary concern to the perceptual system. Object recognition occurs when current perception concurs with stored information about previously encountered objects. Thus, the ventral stream provides the perceptual foundation for the off-line control of action, projecting action into the future, and incorporating stored information from the past into the control of current actions.

Perception, Action, and Consciousness

p0070   As outlined in the section titled 'Different metrics and frames of reference for perception and action,' the ventral and dorsal streams play different but complementary roles in the service of behavior. The ventral stream (together with associated cognitive machinery) permits the brain to identify goals and plan appropriate actions; the dorsal stream (in conjunction with related circuits in premotor cortex, basal ganglia, and brainstem) programs and controls those actions. Ultimately then, both systems transform visual information into motor output. In the dorsal stream, the transformation is quite direct: visual input and motor output are essentially 'isomorphic' with one another. In the ventral stream, however, the transformation is indirect: the construction of a perceptual representation of the visual world enables a more 'propositional' relationship between input and output, taking into account previous knowledge and experience.

p0075   The neuropsychological evidence from patients such as D.F. suggests that there is a close relationship between ventral-stream processing and consciousness. It is not that D.F. is simply unable to describe the form of objects; she seems to have absolutely no conscious appreciation of their dimensions, shape, or orientation. For example, even though D.F.'s hand automatically conforms to the dimensions of an object when she reaches out to grab it, she cannot indicate the width of the object by using an explicit manual 'matching' response, using her index finger and thumb to show how wide she thinks it is. The loss of form perception is only one example of how ventral-stream damage can affect visual consciousness. For example, patients with damage to other ventral-stream areas can lose all conscious experience of color or visual texture. Such patients may continue to see the boundaries or edges between adjoining patches of color, even though they have no appreciation whatsoever of the colors determining those boundaries. Patients with damage to primary visual cortex, which is the major input to the ventral stream, typically report seeing nothing in the visual field contralateral to their lesion. They are cortically blind. But some of these patients show evidence of spared visuomotor control in their blind field, a phenomenon sometimes called blindsight. In other words, they can make eye movements or point to targets that they cannot consciously see, and may even show evidence of anticipatory shaping of the hand when they reach out to grasp objects placed in their blind field. It is thought that blindsight of this kind may rely on projections to phylogenetically ancient visuomotor structures in the midbrain and brainstem, which are in turn connected with the visuomotor systems in the dorsal stream. All of this evidence, as well as a wealth of other data, suggests that the ventral stream is necessary (though not sufficient) for visual consciousness.

p0080   From an evolutionary perspective, the visual phenomenology that arises from ventral-stream processing must confer some kind of advantage to those that possess it. In other words, the capacity for conscious perception must give an organism an edge in natural selection. One can only speculate as to what that advantage might be. It is possible that conscious representations of the world are the only representations that can (eventually) enter our long-term visual knowledge. As already discussed, by retrieving information from long-term memory, we can manipulate this information in working memory (together with information from current percepts) for flexible 'off-line' control of

behavior. Moreover, when an animal is aware of what it sees, it can use this information to decide between different actions (including making no action), to plan future actions, and to communicate what it sees (or has seen) to its conspecifics.

p0085 Being conscious of visual processing in the dorsal stream would confer no such advantages on the animal. Because the transformations that are carried out on visual input in the service of action involve 'just-in-time' computations based on the particular disposition of the goal object with respect to the actor, there would be no value and even a real cost in allowing this information to be accessible to conscious experience. Indeed, if such information were conscious, it would more often than not stand in real contradiction to the scene-based representations of the world offered by the ventral stream, which transcend particular viewpoints and thereby retain their utility over time. Far better then, that dorsal-stream computations should take place as automatically and unconsciously as those of the vestibular system in its efforts to help us maintain an upright and steady posture. In fact, there is now empirical evidence that the visual information that guides our movements is indeed unconscious. Normally, of course, we can consciously see our own arm as we reach toward a target, and we can consciously see other objects in the region of the target that could interfere with our reaching movements. But patients who show 'extinction' (typically following damage to the right parieto-temporal brain region) will often report not seeing a brief stimulus on the left when it is accompanied by a stimulus on the right (i.e., the presence of a stimulus on the right 'extinguishes' the perception of a similar stimulus on the left). Nevertheless, such patients will avoid colliding with an obstacle on the left when reaching toward a goal, even though they report seeing nothing in that part of space.

p0090 Although it is the ventral stream that delivers the contents of our visual consciousness, this does not mean that the workings of the dorsal stream play no role at all in the determining our awareness. For one thing, by virtue of helping to direct our eyes (and shift our attention) between different objects and locations in the environment, the area in the dorsal stream that mediates the visual control of eye movements (and covert shifts of attention) causes changes in the information that can be processed by the ventral stream, and thus the content of our visual awareness. The dorsal stream also participates in consciousness in other a-7.9(co)1i2i.(()()-1(r)-135(ft0(5(e)8l)19.(i)0

stimulus and neural activity in the ventral stream has come from studies using the phenomenon of binocular rivalry, in which different visual stimuli (say a face and a cloudburst pattern) are simultaneously and independently presented to the left and right eye. When human observers are presented with such stimuli, they typically report fluctuations in what they see, sometimes reporting seeing a face and sometime a cloudburst pattern. Only very rarely do they see a 'blend' of the two stimuli. This technique has been used with monkeys by training them to make a manual response to signal a switch between seeing stimulus A (the cloudburst pattern) and stimulus B (the face). In a sense, the monkey was being asked to report on the content of its visual experience. As it turns out, single neurons in the inferior temporal cortex, a higher-order area in the ventral stream, that were tuned to one stimulus (say the face) fired more rapidly when the monkey reported seeing that particular stimulus. This correlation between perceptual report and firing rate was not nearly so strong for neurons in earlier visual areas and was completely absent in primary visual cortex. These striking results demonstrated for the first time that there is a direct link between the activity of particular neurons in the ventral stream and what an animal perceives.

p0110   More recent research with humans shows exactly the same thing. It has now become possible to record from single neurons in humans, typically in patients with epilepsy who have had electrodes implanted in their brains to localize the site where the seizure originates. Using a variant of the binocular rivalry paradigm, researchers have shown that visual neurons in the medial temporal lobe respond only to the perceived stimulus and never respond to the suppressed stimulus presented to the other eye. Similar effects have been observed in functional magnetic resonance imaging (fMRI) studies. For example, when volunteers in the fMRI scanner are presented with rivalrous images of a face and a building to the left and right eye, the activity in two areas in the ventral stream that are differentially selective for faces and scenes fluctuate in a reciprocal fashion with what the volunteers report seeing. Thus, when they report seeing a face, the 'face area' is more active and when they report seeing the building, the 'place' area is more active. The relationship between visual awareness

and brain activity is not seen in the dorsal stream; even when observers are unaware of visual stimuli because of interocular suppression, the fMRI activation elicited by those stimuli is just as great as when those stimuli are consciously perceived.

p0115   Although all these studies (and many others) provide compelling evidence that visual consciousness is closely correlated with activity in the ventral stream, this does not mean that we are conscious of everything that is processed by this system. If this were the case, we would be overwhelmed with information, and consciousness could serve no function. In fact, even in the binocular rivalry experiments, activity was never completely abolished when the monkeys or humans reported not seeing the stimulus – it was just reduced. This may suggest that active inhibitory and/or facilitatory mechanisms are at work – and that for visual awareness to occur activity must exceed some sort of threshold. The neural origins of these inhibitory and/or facilitatory mechanisms are unknown, although regions in prefrontal cortex, medial temporal cortex, and the inferior parietal lobule (regions associated with the switching of visual attention) may play a significant role. But even though ventral-stream areas are subject to such modulatory effects from other brain regions, it is ventral-stream activity itself that determines the content of our visual awareness.

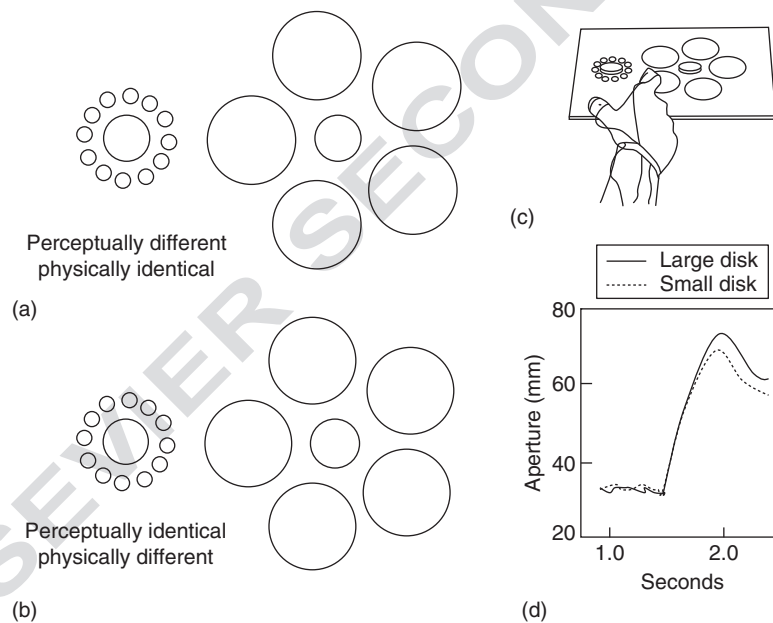## Perception, Action, and Illusions   s0030

p0120   Although much of the evidence for the neural organization of the two visual systems model has come from human neuropsychology and neuroimaging, as well as from work with nonhuman primates, evidence for fundamental differences in the metrics and frames of reference used by vision-for-perception and vision-for-action has also come from studies in normal human observers. These latter studies have also provided compelling evidence with respect to the dissociation between conscious visual experience and more 'automatic' visuomotor control.

p0125   One of the most striking examples of dissociations between vision-for-perception and vision-for-action in normal observers has come from work with pictorial illusions. For example, when

people are asked to pick up a target in the context of a size-contrast illusion, such as the Ebbinghaus Illusion (see **Figure 4**), their grip aperture is typically scaled in flight to the real not the apparent size of the target. Although grip scaling escapes the influence of the illusion, the illusion does affect performance in a manual matching task in which people are asked to open their index finger and thumb to indicate the perceived size of a disk. Thus, the aperture between the finger and thumb is resistant to the illusion when the vision-for-action system is engaged (i.e., when the participant grasps the target) and sensitive to the illusion when the vision-for-perception system is engaged (i.e., when the participant estimates its size).

This dissociation between what people say they see and what they do underscores once more the differences between vision-for-perception and vision-for-action. The obligatory size-contrast effects that give rise to the illusion (in which different elements of the array are compared) normally play a crucial role in scene interpretation and object identification, a central function of vision-for-perception. In contrast, the execution of a goal-directed act, such as manual prehension, requires metrical computations that are centered on the target itself, rather than on the relations between the target and other elements in the scene. In fact, if our visually guided movements were based on the relative rather than the absolute size of objects, then many of our everyday actions, from driving a car to picking up a wine glass, would be subject to critical errors. As it turns out, the true size of a target (for grasping at least) can be computed from the retinal-image size of the object coupled with an accurate estimate of distance based on reliable cues such as vergence of the eyes. Such computations would be quite insensitive to the kinds of pictorial cues that distort perception when familiar illusions are presented.

Figure 4   The effect of a size-contrast illusion on perception and action. (a) The traditional Ebbinghaus illusion in which the central circle in the annulus of larger circles is typically seen as smaller than the central circle in the annulus of smaller circles, even though both central circles are actually the same size. (b) The same display, except that the central circle in the annulus of larger circles has been made slightly larger. As a consequence, the two central circles now appear to be the same size. (c) A 3-D version of the Ebbinghaus illusion. People are instructed to pick up one of the two 3-D disks placed either on the display shown in panel A or the display shown in panel B. (d) Two trials with the display shown in panel B, in which the same person picked up the small disk on one trial and the large disk on another. Even though the two central disks were perceived as being the same size, the grip aperture in flight reflected the real not the apparent size of the disks. Reproduced from Aglioti S, DeSouza JFX, and Goodale MA (1995) Size-contrast illusions deceive the eye but not the hand. *Current Biology* 5: 679–685, with permission from Elsevier.

p0135   There has been considerable debate in the literature about whether or not grasping and other visuomotor responses (such as saccadic eye movements) are refractory to pictorial illusions. In some instances, particularly with unpracticed or awkward movements (where there is a lot of cognitive supervision), grip scaling appears to fall victim to size-contrast illusions like the Ebbinghaus. But of course the fact that actions are sensitive to illusory displays under certain conditions can never by itself refute the idea of two visual systems, which is securely based on a much larger body of evidence ranging from neuroimaging to neurophysiology. Indeed, it is unsurprising that perception affects our motor behavior, even within the context of the two-visual-systems model. After all, there are a number of situations, such as picking up a hammer or a cup of coffee, where our perception of the goal object will determine the kind of grip posture we adopt. The real surprise (at least for monolithic accounts of vision) is that there are a number of situations where visually guided action is unaffected by pictorial illusions. The actions that fall into this category tend to be rapid skilled actions, usually with the preferred hand. Nevertheless, it is fair to say that the claim that actions can be resistant to pictorial illusions is still regarded as controversial, particularly amongst those who favor a more monolithic account of visual processing.

p0140   Most of the pictorial illusions that have been used to dissociate vision-for-perception and vision-for-action distort perceived size by only a few millimeters. There are other illusions, however, which are not only much larger but which also show an actual reversal of depth. One particularly striking example is the hollow face illusion shown in **Figure 5**. In this illusion, knowledge about what faces look like impels observers to see the inside of a mask as if it were a normal protruding face, and the illusory face is perceived to be located several centimeters in front of the actual surface of the hollow mask. Despite the fact that observers cannot resist this compelling illusion, when they are asked to reach out quickly and flick off a small bug-like target stuck on the face, they unhesitatingly reach to the correct point in space (i.e., inside the mask). In other words, despite the presence of a strong hollow-face illusion, people direct rapid movements to the real, not the illusory positions of the targets. To
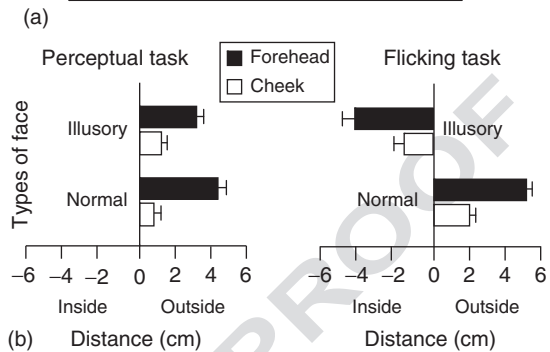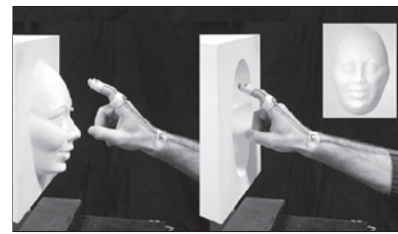

(a)

(b)

Figure 5   Perceptual judgments and visuomotor control with the hollow-face illusion. (a) A small magnet was placed on either the cheek or forehead of the normal face (left) or the hollow mask (right). Participants were required either to flick the magnet from the normal or illusory (actually hollow) face or to estimate its distance psychophysically. Inset shows a photograph of bottom-lit hollow face, in which the illusion of a normal convex face is evident. (b) (Left) The mean psychophysical (perceptual) judgments of the apparent position of the magnets on the illusory and normal face with respect to the reference plate from which the two displays either protruded or receded. Note that participants perceived the hollow face as protruding forward like the normal face. (Right) The mean distance of the hand at the moment the participant attempted to flick the target off the cheek or forehead of the illusory (actually hollow) or the normal face. In the case of the illusory face, the end points of the flicking movements corresponded to the actual distances of the targets, not to consciously seen distances. Error bars indicate the standard error of the mean. Reproduced from Króliczak G, Heard P, Goodale MA, and Gregory RL (2006) Dissociation of perception and action unmasked by the hollow-face illusion. *Brain Research* 1080: 9–16, with permission from Elsevier.

do this, the visuomotor system must have access to a different source of visual information from that driving the illusion. In either case, people seem to be unaware of the veridical depth information they are using to control their flicking movements – and furthermore the use of this information does not 'break' the illusion. Again, this provides compelling evidence that one's conscious perception of a visual

stimulus does not control visuomotor responses directed at that same stimulus.

s0035 ## Time and the Two Streams

p0145 As was discussed earlier, the visuomotor systems in the dorsal stream works in real time, using information provided in a 'bottom up' fashion from the retina. Thus, movements directed to remembered objects (objects that were present, but are no longer visible) might be expected to differ from movements directed to objects in real time. In fact, this is exactly the case. For example, when people reach out to grasp objects that were visible only a few seconds earlier, their grip scaling is now susceptible to pictorial illusions that were present in the display. Such sensitivity is to be expected, of course, if the programming and control of a delayed grasping movement is dependent not on processing in the dorsal stream but on memories derived from perceptual processing in the ventral stream. This conclusion is supported by the fact that the visual agnosic patient, D.F., who has ventral-stream lesions, demonstrates extremely poor size scaling of her grip when she attempts to grasp a target object after it has been removed from view – even though she shows excellent real-time control of her grasping. In one experiment, when a 2 s delay was introduced between viewing the object and initiating the grasp, there was no correlation at all between the size of the object and the aperture of her grasp in flight. An even more surprising result has been obtained in experiments with patients with lesions of the dorsal stream. Even though these individuals have great difficulty scaling their grasp when reaching out to grasp visible objects immediately after the objects are presented, they show a paradoxical improvement in performance if they are required to wait for 5 s before initiating their movement. Again these findings support the idea that the programming and control of delayed actions depends on information derived from earlier perceptual processing in the ventral stream.

s0040 ## Summary

p0150 There appear to be two ways in which visual information can influence and guide behavior. One is immediate and direct. For example, visual information about the size, shape, and disposition of an object with respect to the observer can be automatically transformed into the required egocentric coordinates for the programming and online control of a smoothly executed grasping movement. This kind of visual guidance, which is mediated by the dorsal stream of visual projections, needs to be quick and accurate, and evolution has ensured that it is. The visual information used by the dorsal stream is not accessible to consciousness – even though the actions controlled by that information clearly are. The other way in which vision can influence behavior is much less direct, and depends upon the construction and storage of visual representations that are initially processed in the ventral stream and reflect the structure and semantics of the scene facing the observer. The nature and intentions of subsequent actions will to varying degrees depend on the retrieval, and mental manipulation of, these representations. It is these representations that make up the visual contents of consciousness. The division of labor between the two streams is associated with fundamental differences in the metrics and frames of reference used by vision-for-action and vision-for-perception. Although both streams process information about the structure of objects and about their spatial locations, they use quite different modi operandi to do this. The operations carried out by the ventral stream use scene-based frames of reference and relational metrics; those carried out by the dorsal stream use egocentric frames of reference and absolute metrics. The two streams work together in the production of goal-directed behavior. The ventral stream (together with associated cognitive machinery) identifies goals and plans appropriate actions; the dorsal stream (in conjunction with related circuits in premotor cortex, basal ganglia, and brainstem) programs and controls those actions. Thus, a full understanding of the integrated nature of visually guided behavior will require that we specify the nature of the interactions and information exchange that occurs between the two streams of visual processing.

See also: Neural Basis of Perceptual Awareness (00054); Neuroscience of Volition and Action (00084).

Suggested Readings

Bruno N, Bernardis P, and Gentilucci M (2008) Visually guided pointing, the Müller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: Conclusions from 33 independent studies. *Neuroscience and Biobehavioral Reviews* 32: 423–437.

Carey DP (2001) Do action systems resist visual illusions? *Trends in Cognitive Sciences* 5: 109–113.

Goodale MA (2008) Action without perception in human vision. *Cognitive Neuropsychology* 25: 891–919.

Goodale MA and Milner AD (2004) *Sight Unseen: An Exploration of Conscious and Unconscious Vision*. Oxford: Oxford University Press.

Logothetis NK (1998) Single units and conscious vision. *Philosophical Transactions of the Royal Society of London. B Biological Sciences* 353: 1801–1818.

Milner AD (2008) Visual awareness and human action. In: Weiskrantz L and Davies M (eds.) *Frontiers in Consciousness Research*, pp. 169–214. Oxford: Oxford University Press.

Milner AD and Goodale MA (2006) *The Visual Brain in Action*. 2nd edn. Oxford: Oxford University Press.

Milner AD and Goodale MA (2008) Two visual systems re-viewed. *Neuropsychologia* 46: 774–785.

Biographical Sketch

Professor Goodale was born in England but was educated entirely in Canada. After completing his PhD in psychology at the University of Western Ontario, he returned to the United Kingdom where he worked as a postdoctoral fellow at the University of Oxford with Professor Larry Weiskrantz. After two years at Oxford, Professor Goodale accepted a position in the School of Psychology at the University of St Andrews before returning to Canada in 1977, where he now holds the Canada Research Chair in visual neuroscience at the University of Western Ontario. Professor Goodale is best known for his work on the functional organization of the visual pathways in the cerebral cortex, and was a pioneer in the study of visuomotor control in neurological patients. His recent research uses functional magnetic resonance imaging (fMRI) to look at the activity in the normal human brain as it performs different kinds of visual tasks. Professor Goodale serves on the editorial board of a number of journals including *Experimental Brain Research* and *Neuropsychologia*, and is the past-President of the *Association for the Scientific Study of Consciousness*. He has been recognized for his distinguished scientific achievements by the Canadian Society for Brain, Behaviour, and Cognitive Science, receiving the D.O. Hebb Award in 1999. He was elected as a fellow of the Royal Society of Canada in 2001.

David Milner was born in Leeds, England. After taking his first degree at the University of Oxford, he undertook training in clinical psychology, and then a PhD in behavioral neuroscience, both at the Institute of Psychiatry at the University of London. In 1970, he moved to an academic position at the University of St Andrews, Scotland. It was there that Dr. Milner first collaborated with Dr. Goodale – at that time on research into midbrain visual function in animals. In 1988, he first came across the now-famous patient D.F., who had just suffered catastrophic brain damage due to carbon monoxide poisoning. It was the remarkable early results from testing D.F. that stimulated the development of the idea that the two main visual pathways in the primate brain might play complementary roles in visual perception and visuomotor control, respectively. These ideas were presented in detail in the monograph *The Visual Brain in Action* (Oxford University Press, 1995). In parallel with this work Dr. Milner has been carrying out a program of research into the clinical condition of visuospatial neglect over the past 20 years. David Milner was elected in a fellow of the Royal Society of Edinburgh in 1992, and served as Editor-in-chief of the journal *Neuropsychologia* from 2000 to 2005. He has been director of the cognitive Neuroscience Research unit at Durham University, England, since 2000.