How (and why) the visual control of action differs from visual perception

Melvyn A. Goodale

The Brain and Mind Institute, The University of Western Ontario, London, Ontario, Canada N6A 5B7

Vision not only provides us with detailed knowledge of the world beyond our bodies, but it also guides our actions with respect to objects and events in that world. The computations required for vision-for-perception are quite different from those required for vision-for-action. The former uses relational metrics and scene-based frames of reference while the latter uses absolute metrics and effector-based frames of reference. These competing demands on vision have shaped the organization of the visual pathways in the primate brain, particularly within the visual areas of the cerebral cortex. The ventral 'perceptual' stream, projecting from early visual areas to inferior temporal cortex, helps to construct the rich and detailed visual representations of the world that allow us to identify objects and events, attach meaning and significance to them and establish their causal relations. By contrast, the dorsal 'action' stream, projecting from early visual areas to the posterior parietal cortex, plays a critical role in the real-time control of action, transforming information about the location and disposition of goal objects into the coordinate frames of the effectors being used to perform the action. The idea of two visual systems in a single brain might seem initially counterintuitive. Our visual experience of the world is so compelling that it is hard to believe that some other quite independent visual signal—one that we are unaware of—is guiding our movements. But evidence from a broad range of studies from neuropsychology to neuroimaging has shown that the visual signals that give us our experience of objects and events in the world are not the same ones that control our actions.

1. Introduction

Vision provides our most direct experience of the world beyond our bodies. Although we might hear the wind whistling through the trees and a sudden clap of thunder, vision allows us to appreciate the approaching storm in all its majesty: the roiling clouds, a flash of lightning, the changing shadows on the hillside and the advancing sheets of rain. It is the immediacy of this kind of experience that has made vision the most studied—and perhaps the best understood—of all the human senses.

Much of what we have learned about how vision works has come from psychophysical experiments in which researchers present people with a visual display and simply ask them whether or not they can see a stimulus in the display or tell one visible stimulus from another. By recording the responses of the observers while systematically varying the physical properties of the visual stimuli, researchers have amassed an enormous amount of information about the relationship between visual phenomenology and the physics of light—as well as what goes wrong when different parts of the visual system are damaged.

Vision not only enables us to perceive the world, it also provides exquisite control of the movements we make in that world—from the mundane act of picking up our morning cup of coffee to the expert return of a particularly, well-delivered serve in a game of tennis. Yet, with the notable exception of eye movements, which have been typically regarded as an information-seeking adjunct to visual perception, little attention has been paid to the way in which vision is used to programme and control our actions, particularly the movements of our hands and limbs. One reason why the visual control of action was ignored for so long is the prevalent belief that the same visual representation of the world that enables us to perceive objects and events also guides...
our actions with respect to those objects and events. As I will outline in this paper, however, this common-sense view of how vision works is not correct (for a philosophical discussion of this issue, see [1]). Evidence from a broad range of studies has made it clear that vision is not a unitary system and the control of skilled actions depends on visual processes that are quite separate from those mediating perception.

Of course, it would be convenient for researchers if vision were a single and unified system. If this were the case, then what we have learned from psychophysical studies of visual perception would be entirely generalizable to the visual control of action. But the reality is that visual psychophysics can tell us little about how vision controls our movements. This is because psychophysics depends entirely on conscious report: observers in a psychophysical experiment tell the experimenter about what they are experiencing. But these experiences are not what drive skilled actions. As I discuss below, the control of action depends on visual processes that are quite distinct from those leading to perception, and these processes engage quite different neural circuits. In addition, the visual information controlling action is often inaccessible to conscious report. For all these reasons, conventional psychophysics can provide little insight into the visual control of movement—or tell us anything about the neural substrates of that control. Quite different approaches are needed in which the parameters of visual stimuli are systematically varied and the effects of those variations on the performance of skilled actions are measured. Fortunately, over the last two decades, a new ‘visuomotor psychophysics’ has emerged that is doing exactly this.

Today, researchers are investigating the ways in which vision is used to control a broad range of complex goal-directed actions—and are exploring how the brain mediates that control. New models of the functional organization of the visual pathways in the primate cerebral cortex have been put forward—models that challenge older monolithic views of how the visual system works. In this review, I will trace some of the lines of research that have provided new insights into the role of vision in the programming and control of action and have made the study of vision-for-action a vibrant part of the vision research enterprise.

2. Neuropsychological dissociations between perception and action

In the 1980s, research on the visual abilities of healthy human observers was revealing what appeared to be a striking dissociation between the visual control of skilled actions and psychophysical reports (e.g. [2,3]). By the early 1990s [4], David Milner and I had shown that these dissociations between vision-for-action and vision-for-perception could be mapped onto the two prominent streams of visual projections that arise from the primary visual cortex: a dorsal visual stream projecting to the posterior parietal cortex and a ventral stream projecting to the inferotemporal cortex [5].

According to our two-visual-systems account [4,6,7], the dorsal stream plays a critical role in the real-time control of action, transforming moment-to-moment information about the location and disposition of objects into the coordinate frames of the effectors being used to perform the action. By contrast, the ventral stream (together with associated cognitive networks) constructs the rich and detailed visual representations of the world that allow us to identify objects and events, attach meaning and significance to them and establish their causal relations. Such operations are essential for accumulating and accessing a visual knowledge-base about the world. In addition, the ventral stream provides the perceptual foundation for the off-line control of action, projecting planned actions into the future and incorporating stored information from the past into the control of current actions. By contrast, processing in the dorsal stream does not generate visual percepts; it generates skilled actions (in part by modulating more ancient visuomotor modules in the midbrain and brainstem; see [8]). As one might expect, the two streams are heavily interconnected [9,10]—but at the same time, they play complementary roles in the production of adaptive behaviour. Of course, the dorsal and ventral streams have other roles to play as well. For example, the dorsal stream, together with areas in the ventral stream, plays a role in spatial navigation—and areas in the dorsal stream appear to be involved in some aspects of working memory [9]. In this review, however, I focus on the respective roles of the two streams in perception and action.

So how did we come to propose this division of labour between the two visual streams? Our first insights came from studies of the visual abilities of a neurological patient known by the initials, D.F. [4,11]. D.F. had developed visual form agnosia after being overcome by carbon monoxide from a faulty water heater. Even though D.F.’s ‘low-level’ visual abilities were left relatively intact, she could no longer recognize everyday objects or the faces of her friends and relatives; nor could she identify even the simplest of geometric shapes. It should be emphasized that D.F. has no difficulty identifying an object’s colour or visual texture—and can readily say, for example, whether an object is made of metal, wood or some other substance on the basis of its surface properties [12]. It is the shape of the object that she has problems with. At the same time, she has no trouble identifying the shape of familiar objects by touch. Her deficit in form recognition is entirely restricted to vision.

What is remarkable about D.F., however, is that she shows strikingly accurate guidance of her hand movements when she attempts to pick up the very objects she cannot identify. For example, when she reaches out to grasp objects of different widths, her hand opens wider mid-flight for wider objects than it does for narrower ones, just like it does in people with normal vision [11]. Similarly, she rotates her hand and wrist quite normally when she reaches out to grasp objects in different orientations, and she places her fingers correctly on the surface of objects with different shapes [13]. At the same time, she is quite unable to distinguish among any of these objects when they are presented to her in simple discrimination tests. She even fails in manual ‘matching’ tasks in which she is asked to show how wide an object is by opening her index finger and thumb a corresponding amount. Although some have argued that D.F.’s spared ability to scale her grasping hand to object width depends more on haptic feedback than it does for narrower ones, just like it does in people with normal vision, her deficit in form recognition is entirely restricted to vision. The presence of spared visual control of grasping alongside compromised form discrimination lent additional support to the idea that vision-for-action depends on mechanisms that are quite separate from those mediating vision-for-perception.
But it is the nature of the damage to D.F.’s brain that provided the important clue as to how these different mechanisms might be realized in neural circuitry. An early structural MRI [16] showed evidence of extensive bilateral damage in the ventral visual stream. More recent high-resolution MRI scans [17,18] have confirmed this ventral-stream damage but revealed that the lesions are focused in the lateral occipital cortex (area LO), an area that we now know is involved in the visual recognition of objects, particularly their geometric structure [19]. Even though this selective damage to area LO appears to have disrupted D.F.’s ability to perceive the form of objects, the lesions have not interfered with her ability to use visual form information to shape her hand when she reaches out and grasps objects.

We proposed that D.F.’s spared visually guided grasping is mediated by intact visuomotor mechanisms in her dorsal stream, which, although showing some evidence of cortical thinning at the parieto-occipital junction, appears to be largely intact [18]. Part of the impetus for this idea came from studies showing that patients with damage to the dorsal stream very often exhibit a pattern of deficits and spared abilities—the converse of that observed in D.F. Indeed, it had been known since the early twentieth century that patients with lesions in the posterior parietal cortex can have problems using vision to direct a grasp or aiming movement towards the correct location of a visual target placed in different positions in the visual field, particularly the peripheral visual field [20]. This deficit is often referred to as optic ataxia (following Bálint [20]). But even though these patients show extremely poor visuomotor control, they are often able to judge the relative position of the object in space quite accurately [21,22]. It should be pointed out, of course, that these patients typically have no difficulty using input from other sensory systems, such as proprioception or audition, to guide their movements [20]. Some patients with dorsal-stream damage are unable to use visual information to rotate their hand, scale their grip or configure their fingers properly when reaching out to pick up an object, even though they have no difficulty describing the orientation, size or shape of objects in that part of the visual field.

It was this compelling double dissociation between the effects of damage to the ventral and dorsal streams that initially led us to propose the two-visual-systems model. The model is also supported by a wealth of converging data from neurophysiological studies in monkeys and, more recently, by neuroimaging studies in both neurological patients and healthy people (for reviews, see [6,23,24]). In §3, I highlight some of the neuroimaging work that has looked at activation in D.F.’s brain when she performs perception and action tasks.

3. Neuroimaging studies of the ventral and dorsal streams

With the advent of functional magnetic resonance imaging (fMRI), the number of investigations of the functional organization of the human ventral stream has grown exponentially. Early on, area LO, which is located in the ventrolateral part of the occipito-temporal cortex, was shown to be intimately involved in object recognition (for review, see [19]). As mentioned earlier, D.F. has bilateral lesions that include area LO in both hemispheres. Not surprisingly therefore, an fMRI investigation of activity in D.F.’s brain revealed no differential activation for line drawings of common objects (versus scrambled versions) anywhere in D.F.’s remaining ventral stream, mirroring her poor performance in identifying the objects depicted in the drawings [17]. When healthy control participants were presented with the same line drawings, they exhibited strong bilateral activation of area LO. In fact, when the brains of the control participants were mathematically morphed to align with D.F.’s brain, the form-related activation in these healthy brains coincided almost perfectly with the LO lesions in D.F.’s brain. These neuroimaging findings, which complement the earlier neuropsychological tests of object recognition in D.F. [12], strongly suggest that area LO plays a critical role in form perception.

Although D.F. is unable to discriminate between the widths of different objects because of damage to her ventral stream, her grasping hand is sensitive to an object’s width when she reaches out to pick it up [11]. Neuroimaging of D.F.’s brain has shown that her spared visually guided grasping in D.F. is accompanied by robust activation in a region of the dorsal stream that is typically activated in healthy control participants performing the same kind of task [17]. This area is located in the anterior part of intraparietal sulcus in the same region of the posterior parietal cortex that has been shown to be involved in visually guided grasping in the monkey [25,26] (called AIP in the monkey and hAIP in the human; see figure 1). Importantly, area LO in the ventral stream is not differentially activated when people reach out and grasp objects [26,27], suggesting that this object-recognition area in the ventral stream is not required for the programming and control of visually guided grasping, and that hAIP and other networks in the posterior parietal cortex (in association with premotor and motor areas) can do this independently (figure 1). The idea that hAIP is specialized for the visual control of grasping is supported by the fact that patients with dorsal-stream lesions that are largely restricted to hAIP show deficits in grasping but retain the ability to reach towards objects [28].

Human AIP is one of a number of visuomotor areas that have been identified in the dorsal stream over the last decade [29,30]. An area in the intraparietal sulcus of the posterior parietal cortex has been identified that is activated when participants in the neuroimaging experiment shift their gaze (or their covert attention) to visual targets initially presented in the visual periphery [29,31,32]. As we saw earlier, the lesions associated with the misreaching that defines optic ataxia are typically found in the dorsal stream [21,22]. More recent quantitative analyses [33] of the lesion sites associated with misreaching and other deficits in visually guided movement have identified several key areas in the dorsal stream that map onto areas identified in recent fMRI studies of visually guided reaching [34,35]. Taken together, all this work converges nicely on the results of the earlier neuropsychological studies suggesting that the dorsal stream plays a critical role in the visual control of skilled actions.

It is worth noting that the visual inputs to the visuomotor networks in the dorsal stream do not all originate from V1. It has been known for a long time that humans (and monkeys) with large bilateral lesions of V1 are still capable of performing many visually guided actions despite being otherwise blind with respect to the controlling stimuli (for review, see [6,36,37]). Many of these residual visual abilities, termed ‘blindsight’ by Weiskrantz and co-workers [38], almost certainly depend on projections that bypass the geniculostriate pathway, such as those going from the eye to the superior
colliculus, the interlaminar layers of the dorsal lateral geniculate nucleus or even directly to the pulvinar. There is accumulating evidence that the visual information conveyed by these extra-geniculostriate projections reaches the dorsal stream (for review, see [24]).

4. Different metrics and frames of reference for perception and action

But why are the visual processes underlying visual perception separate from those mediating the visual control of action? The answer to this question lies in the differences in the computational requirements of vision-for-perception on the one hand and vision-for-action on the other. To be able to grasp an object successfully, for example, the visuomotor system has to deal with the actual size of the object, and its orientation and position with respect to the hand used to pick it up. These computations need to reflect the real metrics of the world, or at the very least, make use of learned ‘look-up tables’ that link neurons coding a particular set of sensory inputs with neurons that code the desired state of the limb [39]. The time at which these computations are performed is equally critical. As we move around, the egocentric location of a target object changes from moment-to-moment—and, as consequence, the required coordinates for action need to be computed at the very moment the movements are performed.

By contrast, vision-for-perception does deal with the real-world sizes of objects or their locations with respect to our hands and limbs—a fact that explains why we have no difficulty watching television, a medium in which there are no absolute metrics at all. Our perceptual representations use a scene-based frame of reference, which encodes the size, orientation and location of objects relative to each other. By working with perceptual representations that are object- or scene-based, we are able to maintain the constancies of size, shape, colour, lightness and relative location, across different viewing conditions. The products of perception are available for over a much longer timescale than the visual information used in the control of action. The creation of long-term representations enables us to identify objects when we encounter them again, to attach meaning and significance to them and to establish their causal relations—operations that are essential for accumulating knowledge about the world.

These differences in the metrics, frames of reference and timing of vision-for-action and vision-for-perception can be seen in the behaviour of healthy observers performing tasks that demand either skilled target-directed movements or explicit visual judgements about those the same visual stimuli. Some of the more powerful demonstrations of the dissociation between vision-for-action and vision-for-perception are discussed in §5.

5. Evidence for a perception–action dissociation from behavioural studies in normal observers

Although we have the strong impression that our perception of an object and the visual control of actions we direct at that object depend on one and the same visual representation [1], more than 30 years of work in the laboratory has shown that this is not the case. Some of the earliest and most compelling
evidence for a separation between the processes underlying vision-for-perception and vision-for-action in healthy observers came from experiments in which a target was moved unpredictably to a new location during a rapid eye movement (saccade) directed towards the initial location of the target [3,40,41]. In these experiments, participants typically failed to report the displacement of the target even though a later correction saccade or a manual aiming movement accurately reflected the shift in the target’s position. Even when the participants in these experiments were told that the target would be displaced during the eye movement on some of the trials, they were no better than chance at guessing whether or not such a displacement had occurred. These and other experiments using the same paradigm (for review, see [42]) have consistently shown that participants fail to perceive changes in target position even though they modify their visually guided eye and hand movements to accommodate the new position of the target.

The failure to detect a change in the position of a visual stimulus after a rapid eye movement may reflect a ‘broad tuning’ of perceptual constancy mechanisms that preserve the apparent stability of an object in space as its position is shifted on the retina during an eye movement. Such mechanisms are necessary if our perception of the world is to remain stable as our eyes move about—and the spatial coding of these mechanisms need not be finely tuned. The action systems controlling goal-directed movements, however, need to be much more sensitive to the exact position of a target and, as a consequence, they will compensate almost perfectly for real changes in target position to which the perceptual system may be quite refractory. Thus, because of the different requirements of the perception and action systems, an illusory perceptual constancy of target position can be maintained in the face of large amendments in visuomotor control.

Complementary dissociations between perception and action have also been observed, in which perception of a visual stimulus can be manipulated while a visually guided movement directed at the stimulus remains unaffected. Bridge- man et al. [2], for example, have shown that even though a fixed visual target surrounded by a moving frame appears to drift in a direction opposite to that of the frame, when observers make a rapid aiming movement with their finger, they persist in pointing to the veridical location of the target. In the context of the arguments outlined above, however, such a result is not surprising. Perceptual systems are largely concerned with the relative position of stimuli in the visual array rather than with their absolute position within an egocentric frame of reference. For action systems of course, the absolute position of a visual stimulus is critical, since the relevant effector must be directed to that location. Thus, a perceived change in the position of a target can occur despite the fact that a movement directed at the target is unaffected—and, as we saw in the experiments described above [3,40,41], a ‘real’ change in the position of a target (within an egocentric frame of reference) can produce adjustments to motor output even though that change is not perceived by the participant.

The spatial location of a goal object is not the only visual feature that shows a clear dissociation between perception and action. Our perceptual judgements of object size, for example, are affected by the context in which they are viewed. For example, when we see an average-size person standing next to a professional basketball player, that person can appear to be far shorter than they really are. In the perception literature, there are many examples of size-contrast illusions in which perception of the size of an object does not correspond to its real size. One such illusion is the familiar Ebbinghaus illusion, in which two target circles of equal size are presented, each surrounded by a circular array of either smaller or larger circles. Observers typically report that the target circle surrounded by smaller circles appears larger than the other target circle (figure 2a). Although the perception of size is clearly affected by such contextual manipulations, one might expect that the scaling of grasping movement directed to the central circles might escape this and other size-contrast illusions. After all, in many situations, it is critical that one grasp a target accurately, even when contextual cues might distort one’s perception.

The first demonstration that grasping might be insensitive to the Ebbinghaus illusion was carried out by Aglioti et al. [43]. They constructed a three-dimensional version of the Ebbinghaus illusion, in which a poker-chip type disc was placed in the centre of a two-dimensional annulus made up of either smaller or larger circles (figure 2b). Two versions of the Ebbinghaus display were used. In one case, the traditional illusion, the two central discs were physically identical in size, but one appeared to be larger than the other (figure 2b). In the second case, the size of one of the discs was adjusted so that the two discs were now perceptually identical but had different physical sizes (figure 2c). Despite the fact that the participants in this experiment experienced powerful illusion of size, their anticipatory grip aperture was unaffected by the illusion when they reached out to pick up each of the central discs. In other words, even though their perceptual estimates of the size of the target disc were affected by the presence of the surrounding annulus, maximum grip aperture between the index finger and thumb of the grasping hand, which is typically reached about 70% of the way through the movement, was scaled to the real not the apparent size of the central disc (figure 2d).

The findings of Aglioti et al. [43] have been replicated in a number of other studies (for review, see [24]). Nevertheless, other studies using the Ebbinghaus illusion have shown that action and perception are affected in the same way by size-contrast illusions (e.g. [44,45]), and it has been suggested that the apparent difference in sensitivity between perceptual judgements and grip scaling is due to other factors such as differences in attention and/or the slope of the psychophysical functions describing the relationship between target size and the required response. Nevertheless, studies that have failed to find a dissociation are in the minority. In fact, recent experiments by Stöttinger et al. [46] have shown that even when attentional demands and psychophysical slopes are equated, manual estimates of object size are much more affected by a pictorial illusion (in their case, the Diagonal illusion) than are grasping movements.

But even so, the majority of studies that have claimed that action escapes the effects of pictorial illusions (including the one by Stöttinger et al. [46]) have demonstrated the dissociation by finding a null effect of the illusion on grasping movements. In other words, they have found that perception (by definition) was affected by the illusion, but peak grip aperture of the grasping movement was not. Null effects like this, however, are never as compelling as double dissociations between action and perception. As it turns out, a more recent study has been able to demonstrate a double dissociation between perception and action. Ganel et al. [47]
6. Differences in shape processing for perception and action

The debate about whether or not action escapes the effects of perceptual illusions is far from being resolved (for reviews, see [23,48]). Moreover, the pre-occupation with this issue has directed attention away from the more fundamental question of the nature of the computations underlying visuomotor control. One example of an issue that has received only minimal attention from researchers is the role of object shape in visuomotor control—and how the processing of shape for action might differ from conventional perceptual accounts of shape processing.

The idea that vision treats the shape of an object in a holistic manner has been a basic theme running through theoretical accounts of perception from early Gestalt psychology [49] to more contemporary cognitive neuroscience (e.g. [50,51]). Encoding an object holistically permits a representation of the object that preserves the relationship between object parts and other objects in the visual array without requiring precise information about the absolute size of each of the object’s dimensions. In fact, calculating the exact size, distance and orientation of every aspect of every object in a visual scene would be computationally expensive. Holistic (or configural) processing is much more efficient for constructing perceptual representations of objects. When we interact with an object, however, it is imperative that the visual processes controlling the action take into account the absolute metrics of the most relevant dimension of the object without being influenced by other dimensions or features. In other words, rather than being holistic, the visual processing mediating the control of action should be analytical. Moreover, as discussed earlier, the required coordinates for action could be computed at the moment the action is to be performed: a ‘just-in-time’ computation.

Empirical support for the idea that the visual control of action is analytical, rather than configural comes from experiments in which participants judged the width of a series of rectangular target objects or simply reached out and picked them up across their width [52,53]. The trick was that on some blocks of trials for the length of the objects remained

Figure 2. The effect of a size-contrast illusion on perception and action. Panel (a) shows 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. Panel (b) shows 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. Panel (c) shows the experimental set-up in which the participant was required to reach out and pick up one of two discs that were placed on the illusory display. Panel (d) shows for grip aperture for two trials in which the participant picked up the small disc on one trial and the large disc on another from the display shown in panel (b). Even though the two central discs were perceived as being the same size, the grip aperture in flight reflected the real not the apparent size of the discs. Adapted from Aglioti et al. [43].

used the well-known Ponzo illusion in which the perceived size of an object is affected by its location within pictorial depth cues (figure 3, top). Objects located at the converging end of the display appear to be larger than those located at the converging end. To dissociate the effects of real size from those of illusory size, Ganel and colleagues manipulated the real sizes of two objects that were embedded in a Ponzo display so that the object that was perceived as larger was actually the smaller one of the pair (figure 3, top). When participants were asked to make a perceptual judgement of the size of the objects, their perceptual estimates reflected the illusory Ponzo effect. By contrast, when they picked up the objects, their grip aperture was tuned to their actual size. In short, the difference in their perceptual estimates of size for the two objects, which reflected the apparent difference in size, went in the opposite direction from the difference in their peak grip aperture, which reflected the real difference in size (figure 3, bottom left). This double dissociation between the effects of apparent and real size differences on perception and action, respectively, cannot be explained away by appealing to differences in attention or differences in slope [44,45].
the same from trial to trial (i.e. just the width varied), whereas on other trials both the length and the width varied. As expected, when the length of the rectangular target was varied randomly from trial to trial, participants took longer to discriminate a wide rectangle from a narrow one than when the length did not change. In other words, in this perceptual task, participants showed evidence of configural or holistic processing of the objects: they were unable to ignore length of an object when making judgements about its width. In sharp contrast, however, participants seemed to ignore the length of the target when grasping it across its width. Thus, they took no longer to initiate (or to complete) their grasping movement when the length of the object varied from trial to trial than when its length remained the same. These findings show that the holistic processing that characterizes perceptual processing does not apply to the visual control of skilled actions such as grasping. Instead, the visuomotor mechanisms underlying this behaviour deal with the basic dimensions of objects as independent features.

7. Interactions between the two streams

Although I have been focusing on the division of labour between the dorsal and ventral streams, it is clear that the two streams must interact closely in everyday life. A useful metaphor for understanding the interplay between the ventral and dorsal streams is ‘tele-assistance’, a concept that comes from engineering and robotics. In tele-assistance, a human operator uses a symbolic code to communicate with a semi-autonomous robot that actually performs the required motor act on an identified goal object [54]. Such systems are useful for exploring distant or inhospitable environments, such as the surface of the moon or volcanic craters. In terms of the tele-assistance metaphor, the perceptual system in the ventral stream, with its rich and detailed representations of the world, would be the human operator, whereas the visuomotor system in the dorsal stream is the semi-autonomous robot. Processes in the ventral stream identify a particular goal and flag the relevant object in the scene, perhaps by means of an attention-like process. Once a particular goal object has been flagged, dedicated visuomotor networks in the dorsal stream (in conjunction with related circuits in the motor system) can then be activated to perform the desired motor act. Thus, the object will be processed in parallel by both ventral- and dorsal-stream mechanisms—each transforming the visual information in the array for different purposes. Such simultaneous activation will, of course, provide us with visual experience (via the ventral stream) during the performance of a skilled action (mediated by the dorsal stream).

The two streams interact in other ways as well. Certain objects such as tools, for example, demand that we grasp...
the object in a particular way so that we can use it properly once it is in our hand. Perceptual mechanisms in the ventral stream, which are critical for identifying the tool, presumably enable the selection of the appropriate hand posture for using the tool, whereas the visuomotor mechanisms in the dorsal stream carry out the requisite real-world metrical computations to ensure that the eventual grasp is well-formed and efficient. In this case, both streams would have to interact fairly intimately in mediates the final motor output. Thus, although there is a clear division of labour between the ventral and dorsal streams, the two systems work together in the production of adaptive behaviour.

References

3. Goodale MA, Pelisson D, Prablanc C. 1986 Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. Nature 320, 746 – 750. (doi:10.1038/320746a0)


51. O’Craven KM, Downing PE, Kanwisher N. 1999 fMRI evidence for objects as the units of attentional selection. Nature 401, 584–587. (doi:10.1038/44134)

