

Egocentric Spatial Representation in Action and Perception*

ROBERT BRISCOE

Loyola University New Orleans

Neuropsychological findings used to motivate the “two visual systems” hypothesis have been taken to endanger a pair of widely accepted claims about spatial representation in visual experience. The first is the claim that visual experience represents 3-D space around the perceiver using an egocentric frame of reference. The second is the claim that there is a constitutive link between the spatial contents of visual experience and the perceiver’s bodily actions. In this paper, I carefully assess three main sources of evidence for the two visual systems hypothesis and argue that the best interpretation of the evidence is in fact consistent with both claims. I conclude with some brief remarks on the relation between visual consciousness and rational agency.

1. Introduction

The purpose of this paper is to defend a pair of widely accepted claims about spatial representation in visual perception. The first is the claim that visual experience represents 3-D space around the perceiver using a perspectival, egocentric frame of reference. The second is the claim that there is a constitutive link between the spatial contents of visual experience and the perceiver’s bodily actions. Variants of both claims have played a significant role in recent philosophical reflection on perception (Peacocke 1992, chap. 3; Campbell 1994; Brewer 1997, chap. 6; Grush 1998, 2000; Hurley 1998; Mandik 1999; Cussins 2003; Kelly 2004; Noë 2004; Gallagher 2005; Millikan 2006). Nonetheless, proponents of the “two visual systems” hypothesis (Milner & Goodale 1995/2006, Goodale & Milner 2004) have maintained on empirical grounds that both claims are false. In what follows, I carefully review and assess three main sources of evidence for the two visual systems hypothesis (TV): neuropsychological demonstrations that brain damage can have different and separate effects on visual awareness and visual control of action; behavioral studies of normal subjects involving visual illusions; and, last, speculation about the computational demands made by conscious seeing, on the one hand, and fine-grained, visually based engagement with objects, on the other. Contrary to well-

* For helpful discussion of earlier versions of this paper and/or encouragement I am grateful to Avner Baz, Joe Berendzen, Juliet Floyd, Aaron Garrett, Ruth Garrett Millikan, Alva Noë, Dasha Polzik, Thomas Sattig, and an audience at the Boston Colloquium for Philosophy of Science in January 2006. Both Larry Hardesty and an anonymous referee provided detailed comments and suggestions that significantly improved the final version. Thanks also to Daniel Dennett for conversations about J.J. Gibson that provided the initial impetus to write this paper.

known assessments by Andy Clark (1999, 2001) and John Campbell (2002), I argue that the best interpretation of the evidence in each case is actually consistent with both of the aforementioned claims.

The plan for this paper is as follows: In §2, drawing inspiration from the writings of Gareth Evans, I present what I call the “Egocentric Conception” of the spatial representational content of visual experience (EC).¹ According to EC, personal-level visual experience represents the world using the same egocentric spatial coding system as is used in proprioception to represent the current configuration and orientation of one’s body and, crucially, as is used in planning one’s intentional movements and actions. In §3, I distinguish EC from what Clark (2001) calls the “Assumption of Experience-Based Control” (EBC) and from what Campbell (2002) calls the “Grounding Thesis.” Since both of the latter views are plausibly threatened by TV it is important to show that their rejection is compatible with adherence to EC. In §§4-8, I review and assess the empirical evidence marshaled by Milner and Goodale for TV. I argue that the best interpretation of the evidence in each case is consistent with EC. I also try to show that an EC-friendly interpretation of findings concerning the comparative effects of size-contrast illusions on visual awareness and visuomotor action avoids certain serious theoretical difficulties faced by TV. In §9, I conclude with a couple of brief remarks on the relation between perceptual consciousness and rational agency.

2. Evans on Egocentric Spatial Representation

In *The Varieties of Reference* (1982) and in his paper “Molyneux’s Question” (1985), Gareth Evans argues that in order to specify the spatial information conveyed by a visual experience to its subject it is necessary to use “egocentric terms... that derive their meanings in part from their complicated connections with the subject’s *actions*” (1982, 155). Evans’s proposal comprises two claims. The first is the claim that visual experience represents 3-D space around the subject using a perspectival, egocentric frame of reference. To see a matchbox as over there, e.g., is perforce to perceive it as located somewhere relative to *here*, somewhere, that is, more precisely specified using the axes *right/left*, *front/behind*, and *above/below*. It is to perceive the matchbox as occupying a region of visible space specified in relation to the current location and orientation of one’s own body. Christopher Peacocke’s (1992, chap. 3) proposal that the representational content of a visual experience is given by a spatial type he calls a “scenario” is one familiar elaboration of this view. Individuating a scenario involves specifying which scenes – which ways of “filling out” space around the perceiver at the time of the experience – are consistent with the content’s correctness. Each such scene is constituted by an assignment of surfaces and surface properties (orientations, textures, colors, etc.) to points in a spatial coordinate system whose axes originate from the center of the perceiver’s chest.

Talk of an egocentric frame of reference need not be taken to imply that visual perception organizes the spatial layout of visible objects and surfaces around a single, privileged bodily origin, e.g., a point in (or on) the perceiver’s torso (as in the framework Peacocke develops), or the perceiver’s center of gravity, or the apex of the solid angle of the perceiver’s visual field. Indeed, there is ample evidence

¹ By the “spatial representational content” of a visual experience I intend the condition(s) under which the spatial information conveyed by the experience to its subject would be correct. For discussion of the varieties of representational content (and corresponding varieties of representational correctness), see Rescorla 2003; Millikan 2004, chap. 7; and Heck *forthcoming*.

from cognitive neuroscience that visual systems in the brain construct multiple representations of 3-D space using a variety of coordinated, effector-specific frames of reference (Paillard 1991, Rizzolatti *et al.* 1997, Colby 1998, Colby & Goldberg 1999, Cohen & Andersen 2002). Eye-centered, head-centered, and torso-centered frames of reference, e.g., are coordinated by calibrating continuously updated proprioceptive information about the eye's orientation relative to the head, and the head's orientation relative to the torso. Given information about the location of an object in a torso-centered frame of reference, proprioceptive information about the angles of the perceiver's wrist, elbow, and shoulder joints can then be used to establish its location relative to her hand. Similar transformations are possible for all other coordinated, effector-specific frames of reference.²

This subpersonal representational arrangement seems to be reflected at the personal level. When I see an object's egocentric location, I do not simply see its location relative to me. Indeed, there is no privileged point in (or on) my body that counts as *me* for purposes of characterizing my perceived spatial relation to the object. Rather, my visual experience of the object conveys information about the object's location relative to *every* part of my body (seen or unseen) of which I am proprioceptively aware.³ I perceive, e.g., that a certain book is closer to my right hand than to my left hand, above my waist, but below my chin, etc. Such perspectival, body-relative spatial information – which may be more or less *precise*, depending inter alia on the relevant effector (eye, head, hand, etc.), the object's distance in depth (Cutting & Vishton 1995), and the visual structure of the object's background (Dassonville & Bala 2004), and which may be more or less *salient*, depending inter alia on the specific task situation, the perceiver's expertise, and correlative demands on her attention – is part of the content of a visual experience of an object and is reflected in its phenomenology. Thus, as Peacocke observes, the experience of seeing Buckingham Palace while looking straight ahead is not the same as the experience of seeing the Palace with one's face turned toward it, but with one's body turned to the right (1992, 62). The field of view in each case is identical, but the position of the Palace with respect to the mid-line of one's torso (and, so, the direction of egocentric "straight ahead") is different. Full discussion of the role played by somatic proprioception in the phenomenology of perceptual experience is clearly beyond the scope of this paper.⁴ For present purposes, it is enough to emphasize that the axes in relation to which object positions are represented in visual experience may be presumed to originate from a variety of bodily loci. Crucially, the different effector-specific frames of reference defined by these axes are coordinated using continuously updated proprioceptive information about the spatial configuration of the body. Accordingly, in what follows, I shall understand localization of an object in an egocentric frame of reference to entrain its localization in multiple, coordinated, effector-specific frames of reference.

² For discussion and philosophical applications, see Grush 2000. I should note that there is experimental evidence that the brain may make use of more than one type of egocentric spatial coding system. See Pesaran *et al.* 2006 for evidence that some neuron populations in dorsal premotor cortex utilize a "relative position code" to represent the configuration of the eye, hand, and visual target. In such a coding system, the same pattern of neuronal response is observed whenever the eye, hand, and target occupy the same relative positions regardless of their absolute positions in space.

³ As Anthony Marcel writes, "the 'ego' of an 'egocentric' reference frame is whatever is the body part, and its indexical location, that stands in an intentional relation to a target, and this can be multiple points simultaneously" (2003, 84).

⁴ But see O'Shaughnessy 1980, Bermúdez 1998, Gallagher & Marcel 1999, Grush 2000, Roessler & Eilan 2003, and Gallagher 2005.

Evans's second claim is that the egocentric terms used to specify the spatial information conveyed by a visual experience to its subject derive their meanings from their complicated connections with the subject's *actions*. He writes:

Egocentric spatial terms are the terms in which the content of our spatial experiences would be formulated, and those in which our immediate behavioural plans would be expressed. This duality is no coincidence: an egocentric space can exist only for an animal in which a complex network of connections exists between perceptual input and behavioral output. A perceptual input... cannot have spatial significance for an organism except in so far as it has a place in such a complex network of input-output connections (1982, 154).

To perceive a teapot's egocentric location, on this view, is to acquire an implicit, practical understanding of which direction one would have reason to move, or point, or turn, were it one's intention to move, or point, or turn in the direction of the teapot, and so on, for all such spatially directed actions. Similarly, to perceive the teapot's 3-D shape is to acquire a practical understanding, among other things, of which movements of one's arms and hands would one enable one to touch or to take hold of the teapot. One is conscious, e.g., that one would need to move the tip of one's index finger a bit further and over to the left in order to make contact with *this* (egocentrically identified) point on the handle as opposed to *that* (egocentrically identified) point on the spout.⁵ In general, one is perceptually aware of the region occupied by an object in egocentric space to the extent that one has a practical understanding of the way one's bodily spatial relations to the object would vary with possible intentional movements and actions. Closely related views are familiar from the writings of Merleau-Ponty (1945/1962) and J.J. Gibson (1966, 1979).

I have suggested in the foregoing that both visual perception and visuomotor action make use of an egocentric frame of reference. Although space does not permit adequate elaboration, I should point out that it is plausible that the bodily space of proprioception is also an egocentric space in the sense defined here.⁶ Like visual perception, proprioception makes use of a set of coordinated, effector-specific frames of reference. It does not encode the locations and movements of the subject's limbs in relation to a single, privileged bodily origin. For this reason, it does not make sense to think, e.g., of my left hand as being proprioceptively represented as nearer or further away from *me* than my right foot (Bermúdez 1998, 153). My distinctive, proprioceptive awareness of my hand's location is not an awareness of its location relative to a single, privileged bodily origin, but rather an awareness of its location relative to the various other proprioceptively represented

⁵ It seems clear that one is no more aware of all the low-level details of such potential motor actions – the precise contractions of muscles, adjustments of joint angles, etc. – than one is aware of all the low-level details of one's *actual* motor actions. For discussion of limitations on motoric awareness, see Jeannerod 1997, Blakemore *et al.* 2002, Gallagher & Jeannerod 2002, and Marcel 2003.

⁶ One vivid piece of experimental evidence for this view is provided by the classic "Alien-Hand" experiment and variations thereon (Nielsen 1963, Sørensen 2005). In the experiment, illusory, egocentrically coded *visual feedback* about the movements of the subject's hand in a line-drawing task dominates and significantly distorts the subject's proprioceptive awareness of the hand's movements. Such multisensory integration of spatial information about bodily disposition strongly suggests that perception and proprioception make use of a common, egocentric coding system.

parts of my body (my eyes, head, torso, etc.).⁷ As Brian O’Shaughnessy writes, “the basic ‘given’ is, not just feeling, not just feeling-in-a-certain-body-part, but feeling-in-a-certain-body-part-at-a-position-in-body-relative-physical-space; and so, also, certain-body-part-at-a-position-in-body-relative-physical-space” (1980, 165). For further philosophical defense of this view, see Peacocke 1992, chap. 3; Brewer 1995; and Gallagher & Marcel 1999.

Integral to the “Egocentric Conception” of the spatial representational content of visual experience (EC), as I shall call it, then, is the view that spatial properties are represented in visual experience using the same egocentric spatial coding system, i.e., using the same coordinated, effector-specific frames of reference, as are used in proprioception to represent the current configuration and orientation of one’s body and, crucially, as are used in planning one’s intentional movements and actions. One notable upshot of this action-oriented, unified coding view is that there is no general problem about how the spatially contentful deliverances of visual experience are able to bear upon our bodily motor engagements with the world. The testimony of the senses is delivered in an egocentric “language” that the body understands. Hence, the spatial contents of visual experience can be immediately imported into the contents of intentions for spatially directed, bodily action (Peacocke 1992, chap. 3).

This view seems to comport nicely with the phenomenology of visuomotor action. When I see a fork at a certain egocentric location in front of me on the dinner table, e.g., my visually based awareness of the fork’s position relative to my right hand (even if my hand is hidden from sight at the time of the experience) is typically as direct and authoritative as my proprioceptive awareness of my hand’s position relative to the rest of my body. Hence, no conscious ratiocination or reflection is typically required in order to effect the movements of my hand and arm necessary in order to take hold of the fork by its handle. The information about the object’s coordinates relative to the current spatial configuration of my body, i.e., the egocentric information needed for purposes of initiating bodily movements accurately targeted on the object, is already part of the content of my conscious visual experience. I take it that similar considerations motivate Sean Kelly (2004) to claim that a subject’s visual experience of an object crucially comprises a “bodily readiness” to deal with the object’s spatial features.

3. The Assumption of Experience-Based Control and the Grounding Thesis

EC, as characterized here, claims that there is a constitutive connection between spatially contentful visual experience and visuomotor action. In this respect, it seems clear that EC – or a conception closely akin to EC – is implicit in much philosophical theorizing about visual representation of spatial properties. Recently, however, some philosophers, including Andy Clark (1999, 2001) and John Campbell (2002), have argued that this claim is compromised by an array of empirical evidence for what David Milner and Melvyn Goodale call the “two visual systems hypothesis” (TV). Indeed, in Clark’s view, the evidence suggests that there is “a deep and abiding dissociation between the contents of conscious seeing, on the one hand, and the [representational] resources used for the on-line guidance of visuomotor action, on the other” (Clark 2001, 495).

According to Clark (2001), philosophical reflection on the relation between

⁷ Hence, the finding that proprioceptive illusions with respect to the orientation of one’s head induce correlative proprioceptive illusions with respect to the orientation of one’s arm (Knox *et al.* 2005).

action and perception is frequently premised on what he calls the “Assumption of Experience-Based Control” (EBC):

Conscious visual experience presents the world to the subject in a richly textured way; a way that presents fine detail (detail that may, perhaps, exceed our conceptual or propositional grasp) and that is, in virtue of this richness, especially apt for, and typically utilized in, the control and guidance of fine-tuned, real-world activity (2001, 496).

EBC, as characterized by Clark, has an industrial strength counterpart in what Campbell (2002) calls the “Grounding Thesis.” According to the Grounding Thesis, the spatial parameters (motor coordinates) for one’s visually based action on an object are fully determined by and, in this sense, “grounded” in aspects of one’s conscious perceptual experience of the object. When one reaches for an object, Campbell writes, “the visual information that is being used in setting the parameters for action *must* be part of the content of [one’s] experience of the object” (2002, 50, my emphasis).

It seems clear that the Grounding Thesis is much stronger than EBC. The Grounding Thesis maintains that it is the proprietary functional role of conscious visual experience to control and guide visually based action and, so, that conscious visual experience is *necessary* for visually based action. EBC is less demanding. It maintains only that conscious visual experience is typically utilized in visually based action. Hence, it is able to allow that visuomotor systems may sometimes make use of nonconscious visual spatial information.

Clark argues that the empirical findings and theoretical considerations Milner and Goodale marshal to motivate TV, to be reviewed below, challenge EBC. They challenge the idea that conscious visual experience presents the world to the subject in a way that is both especially apt for and typically utilized in on-line visuomotor action. Similarly, Campbell argues that the array of putative evidence for TV challenges the Grounding Thesis.⁸ It challenges “the idea that it is experience of the location of the object that causally controls the spatial organization of your action on the thing” (2002, 51).

The first remark I should like to make is that neither EBC nor the Grounding Thesis has much *prima facie* plausibility. One reason is that there are, in fact, many familiar examples of visually transduced information subserving finely tuned action in the absence of conscious seeing. In navigating a busy city sidewalk while conversing with a friend, or returning a fast tennis serve,⁹ or driving a car while deeply absorbed in thought, one’s bodily responses and adjustments often seem to be prompted and guided by the nonconscious use of visual information. Many other familiar cases of attentionally recessive visuomotor control could, of course,

⁸ Clark suggests that both teleological accounts of conscious perceptual content such as Millikan 1984 and functionalist accounts such as Dennett 1991 are at risk if EBC is false (2001, 499). It is disputable, however, that either Millikan or Dennett actually relies on EBC. See, e.g., Millikan 2004, chap. 1 and Dennett 1991, chap. 6 for considerations that seem to threaten EBC.

⁹ This example is due to Clark 2001. A related example comes from studies of visual attention in ping-pong and cricket. Researchers have found that subjects nonconsciously utilize visual information in making rapid eye movements that accurately anticipate the ball’s bounce point several hundred milliseconds before the ball actually reaches it (Land & Furneaux 1997, Land & McLeod 2000). For discussion of the pervasive involvement of “zombie agents” in everyday sensorimotor activities, see Gazzaniga 1998 and Koch 2004, chaps. 12 and 13.

be adduced. The involvement of “nonconscious, fine-action guiding systems,” as Clark himself observes, “is sufficiently immense and pervasive, in fact, as to render initially puzzling the functional role of conscious vision itself” (2001, 509). At any rate, it is sufficiently immense and pervasive as to render both EBC and the Grounding Thesis as untenable. Another reason to regard the Grounding Thesis, in particular, as implausible is that, as Evans writes “it seems abundantly clear that the evolution could throw up an organism in which such advantageous links [between sensory input and behavioral output] were established long before it had provided us with a conscious subject of experience” (1985, 387).¹⁰ But, if this is the case, then clearly conscious visual experience cannot be necessary for all environmentally responsive, visually based action.

For present purposes, it is important to emphasize that EC is not premised on either EBC or the Grounding Thesis. EC is a conception of the spatial contents of personal-level visual awareness. It does not make any pronouncements about the extent to which visually based action is possible *without* visual awareness. Hence, in contrast with the Grounding Thesis, EC does not claim that visual awareness of space is implicated in *all* visually based action. (Rather, it claims the converse, that the possibility of intentional, visually based action is implicated in all visual awareness of space.) And, hence, in contrast with EBC, EC does not claim that visual awareness of space is *typically* implicated in visually based action. EC is compatible with the evidence adduced above that, in a wide range of cases, motor systems may make use of nonconscious visuospatial information.

It follows, contrary to Clark, that EBC is not “implicated in any philosophical account which seeks to fix the contents of perceptual experience (whether conceptualized or not) by invoking direct links with action” (2001, 499). Indeed, neither EBC nor the Grounding Thesis, we have just seen, is implicated in EC. An important question, however, remains: Does the evidence for TV challenge EC? In particular, does it challenge either the idea that visual experience represents 3-D space around the perceiver using an *egocentric* frame of reference (Evans’s first claim) or the idea that there is a constitutive connection between the spatial contentfulness of visual experience and the perceiver’s *bodily actions* (Evans’s second claim)? It is to this question that I now turn.

4. The Two Visual Systems Hypothesis

According to TV, the primate brain comprises two, functionally dissociable visual systems: a phylogenetically ancient system subserving visually based action and a phylogenetically recent system subserving conscious visual awareness (figure 1). The former system is identified with the putative *dorsal* processing stream from primary visual cortex (V1) to posterior parietal cortex, while the latter system is identified with the putative *ventral* processing stream from primary visual cortex to inferotemporal cortex. The hypothesized “action” and “perception” systems can be succinctly distinguished as follows:

Action: The action system is concerned with the control and guidance of visually based actions. It contains an array of dedicated visuomotor modules that transform visual inputs into spatially directed motor outputs. Dorsal processing supporting the action system codes fine-grained metrical

¹⁰ Evans cites the case of blindsight famously described in Weiskrantz *et al.* 1974 to illustrate this possibility.

information about the *absolute* size, distance, and geometry of objects in an egocentric frame of reference. Upstream, i.e., “bottom-up,” sources of 3-D spatial information to dorsal processing are quite limited. These include stereopsis (binocular disparity), convergence, and motion parallax.¹¹

Perception: The perception system subserves conscious, high-level recognition of objects and their task-relative significance or function. It is also implicated in the selection of targets for the visuomotor system, e.g., a hatchet, and in the selection of *object-appropriate* types of action in which to engage, e.g., taking hold of the hatchet by its handle. Crucially, ventral stream processing supporting the perception system codes only coarse-grained metrical information about the *relative* size, distance, and geometry of objects in an *allocentric* or *scene-based* frame of reference. Upstream sources of 3-D spatial information to ventral processing are quite extensive. In addition to binocular cues such as stereopsis and convergence, these include monocular, “pictorial” cues such as occlusion, relative size, shading, and reflections, as well as gestalt principles of perceptual organization. Downstream, i.e., “top-down,” sources of 3-D spatial information include stored knowledge about specific types of objects and scenes.

In representing spatial properties, the two hypothesized visual systems are thus taken to contrast significantly in respect of the metrics, the frames of reference, and the sources of spatial information they respectively exploit.¹²

¹¹ As we shall see (§§6-7), however, additional sources of 3-D spatial information are available to the dorsal stream through cross connections with the ventral stream.

¹² Matters are complicated by three observations. First, there are multiple *anatomical* connections between the two putative processing streams, and there is evidence that they may substantially interact via early visual areas such as primary visual cortex. As Goodale & Milner 2004 note, feedback pathways to lower-order areas in the visual brain frequently outnumber feed-forward pathways to higher-order areas (102). Second, there are strong reasons to think that substantial interaction between the two streams is *functionally* necessary for a wide variety of familiar actions. The movements one makes in picking up a cup of coffee, e.g., are determined not only by its visible spatial properties – its shape, location, etc. – but also by its weight, how full the cup is, and by the temperature of the coffee (for related examples, see Peacocke 1993, Jeannerod 1997, Jacob & Jeannerod 2003, and Glover 2004). Plausibly, the mechanics of spatially directed actions involving high-level, stored object knowledge would be determined by both dorsal and ventral processing areas in normal subjects. In tandem, these two observations suggest that the story of the relation between the two streams may be that of mythical Alpheus and Arethusa writ large. (This conclusion is further reinforced by evidence reviewed in §§6-7 below that the dorsal stream accesses or “taps” depth-specific information in the ventral stream whenever it can afford to do so.)

The third observation is that certain dorsal areas seem to subservise both action *and* visual awareness of objects in space (Rizzolatti & Matelli 2003, Gallese 2007). There is a strong correlation, as pointed out below, between damage to the superior part of the posterior parietal lobe and visuospatial attentional deficits. Subjects with such damage, e.g., may sometimes exhibit visual “extinction” in which stimuli on the contralesional side of visual space are sometimes completely ignored, i.e., absent from visual experience. This suggests that visuospatial awareness cannot be neatly mapped onto ventral processing areas.

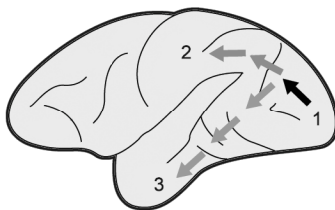


Fig. 1. A sideways view of the macaque monkey brain. Dorsal processing stream from primary visual cortex (1) to posterior parietal cortex (2). Ventral processing stream from primary visual cortex (1) to inferotemporal cortex (3).

Milner and Goodale explain the relationship between the two systems by analogy with the relationship between a human operator and a semiautonomous robot guided by *tele-assistance* (1995/2006, 231-34; 2004, 98-101). In tele-assistance, a remote human operator identifies a goal object, flags the target for the robot, and specifies an action on the target for the robot to perform. Once the relevant information has been communicated to the robot, the robot uses its own sensing devices and processors to determine which movements would enable it to achieve the remotely specified goal. Campbell uses a similar analogy in order to explain the relationship between conscious seeing and visually based action:

There is an obvious analogy with the behaviour of a heat-seeking missile. Once the thing is launched, it sets the parameters for action on its target in its own way; but to have it reach the target you want, you have to have it pointed in the right direction before it begins, so that it has actually locked on to the intended target (2002, 56).

Notably, both analogies assume that the target-selecting system (the ventral stream) has no difficulty in communicating to the target-engaging system (the dorsal stream) with which object it is to interact. This assumption, however, is quite substantial in view of the consideration that, according to TV, the two systems are locating objects using fundamentally different spatial frames of reference. (I shall return to this point in §7 below.)

Evidence for TV comes from three main sources: neuropsychological demonstrations that *brain damage* can have different and separate effects on visual awareness and visual control of action; behavioral studies of normal subjects involving *visual illusions*; and, last, speculation about the *computational demands* respectively made by conscious seeing, on the hand, and fine-grained, visually based engagement with objects, on the other. In what follows, I shall carefully review and assess each source of evidence in turn.

5. Profound Visual Form Agnosia and Optic Ataxia

In normal experience, visual awareness and visually based action function harmoniously. As Wolfgang Prinz, remarking on the typically seamless integration of spatial vision and action, writes, “[a] person, in the analysis of his/her conscious experience, would not find any indication of a qualitative difference or even a gap between action-related and perception-related contents.... Rather, there is a clear

sense of acting in the same world as is perceived" (1990). Normal visual awareness and normal visuomotor skills, however, can come apart in consequence of severe brain damage. Thus a subject, DF, suffering from profound *visual form agnosia* due to a ventral stream lesion, cannot consciously see a target's shape, size, orientation, or location. DF cannot tell whether she is viewing a circle or a triangle or whether a pencil held in front of her is vertical or horizontal. Nor is she able to make copies of simple line drawings. She is well able, however, to make precise, target-directed movements. She can retrieve target objects, scaling her grip aperture to the size of the object, and she can place a card through a slot, rotating her hand to the correct orientation as she extends her arm. She is even able to catch objects that are thrown to her and to walk through test environments while avoiding block obstacles as confidently as control subjects. Nonetheless, if the story Milner and Goodale tell is to be believed, DF is unable to make even simple visual judgments about the spatial layout of her surroundings. She is, one might say, "space blind."

Optic ataxia, caused by damage to superior parietal areas in the dorsal stream, by contrast, does not impair visual acuity or visual attentional abilities, but it does impair visual control of hand and arm movements directed at objects in the periphery of the contralesional visual field (Perenin & Vighetto 1988). For example, optic ataxics may miss target objects by a few centimeters when reaching in their direction, and they show poor scaling of grip and hand orientation when attempting to take hold of them, especially when using the contralesional hand.¹³ Further, in reaching tasks, optic ataxics have difficulty in making fast corrective movements to a jump in target location (Battaglia-Mayer & Caminiti 2002). That said, pure optic ataxics with unilateral lesions are in general able accurately to perform hand and arm movements targeted on stationary objects in central (foveal) vision (Rossetti *et al.* 2003).

How do these findings comport with the idea – integral to EC – that there is a constitutive link between spatially contentful visual awareness and visually based action? Profound visual form agnosia is actually the easier case for EC. Since EC is not premised on the Grounding Thesis, i.e., since EC does not claim that it is the proprietary functional role of conscious visual experience to guide visually based action, it can allow that visually transduced information may, in a variety of cases, subserve environmentally responsive behavior *without* visual awareness. So the possibility of profound visual form agnosia does not by itself seem to present any sort of general, empirical challenge to EC.

Turning now to optic ataxia, the main consideration is that optic ataxia does not involve anything like a complete dissociation of perception and action. In pure cases of optic ataxia, action is impaired only with respect to the precision and fluency with which certain visually guided reaching and grasping movements are performed (hence 'ataxia'). Otherwise, subjects retain fully normal visual control of eye movements, head movements, locomotion, etc. The visuomotor deficits associated with optic ataxia, in other words, are *partial* and *effector-specific*.¹⁴ Indeed, for this reason, optic ataxia provides compelling evidence for the view that "there

¹³ For recent experimental evidence that optic ataxia may be caused in part by a proprioceptive deficit with respect to the location of the contralesional hand, see Blangero *et al.* 2007. In this study, researchers found that optic ataxics make significant errors (relative to control subjects) when pointing in the dark to the location of their contralesional (ataxic) hand using their ipsilesional (normal) hand and vice versa.

¹⁴ As noted above, pure optic ataxics with unilateral lesions are able accurately to grasp stationary objects presented in central vision. They are also able to draw "quite recognizable copies" of pictures they are shown (Goodale & Milner 2004, 36). These considerations indicate that the relevant deficits are partial even in respect of visually guided hand movements.

are multiple spatial codings, each controlling a different effector system" (Milner & Goodale 1995/2006, 94).

Optic ataxia may well indicate a pronounced modularity in the brain in respect of visually guided reaching and grasping, then, but it does not by itself seem to pose a general challenge to the notion that there is a constitutive connection between spatially contentful visual awareness and visually guided action. Optic ataxia evidences only the possibility of normal spatial awareness in the face of partial and effector-specific visuomotor breakdown. It does not evidence the possibility of a complete dissociation of perception and action, i.e., normal spatial awareness in the absence of all abilities to orient toward stimuli, move toward stimuli, track stimuli, etc. So the possibility of optic ataxia does not by itself seem to present any sort of general, empirical challenge to EC. Indeed, normal spatial awareness is presumably possible in optic ataxia because effector systems involved in selective visual attention are left intact. When these systems are pathologically affected, spatial awareness is far from normal. Thus Balint syndrome, also caused by damage to the superior parietal lobe, is characterized not only by optic ataxia, but also by "sticky" or paralyzed gaze (Moreaud 2003). In consequence, a subject with Balint syndrome may have great difficulty in perceiving more than one object or part of space at a time. Damage to the superior parietal lobe may also sometimes result in the phenomenon known as "visual extinction." Extinction patients are able to perceive a visual stimulus as long as it is presented in isolation. When two stimuli are presented at the same time, however, the stimulus on the more contralesional side of visual space is completely ignored, i.e., is absent from visual experience. These deficits suggest that while an intact dorsal stream may not be functionally sufficient for normal visuospatial awareness (as evidenced in part by the profound visual form agnosia consequent upon trauma to the ventral stream in subject DF), much dorsal stream processing may nonetheless be functionally necessary. For further evidence in support of this view, see Rizzolatti & Matelli 2003 and Gallese 2005, 2007.

6. The Argument from Illusion Studies

A second – and much more controversial – source of evidence for TV come from behavioral studies of the comparative effects of size-contrast illusions on visual perception and visuomotor action. Milner & Goodale 1995/2006, chap. 6 and Goodale & Milner 2004, chap. 6 appeal to an experiment conducted by Aglioti *et al.* 1995 involving the Titchener (Ebbinghaus) Circles illusion. In figure 2a, the two central circles perceptually appear to be different in size although they are physically identical, while in figure 2b the two central circles perceptually appear to be identical in size although they are physically different. (In 2b, the central circle on the right has been enlarged in order to appear as the same size as the central circle on the left.) The illusion presumably arises due to the way the visual system is influenced by the contextual, pictorial depth cues in the display, in particular, the relative size of the circles.¹⁵ In 2a, the circles in the annulus (ring) around the central circle on the left are much smaller than the circles in the annulus around the central circle on the right and, so, are interpreted – together with the central circle on the left – as more distant in depth. But because the two central circles in 2a are

¹⁵ Relative-size is a "contextual" or "context-dependent" pictorial depth cue because it involves comparisons between different objects in the scene. In addition to relative size, other contextual, pictorial depth cues include relative density, occlusion, height in the visual field, and aerial perspective. See Vishton & Cutting 1995 and Palmer 1999, chap. 5 for discussion.

physically the same size (and, so, subtend the same visual angle), the central circle on the left is perspectively interpreted as being larger than the central circle on the right.

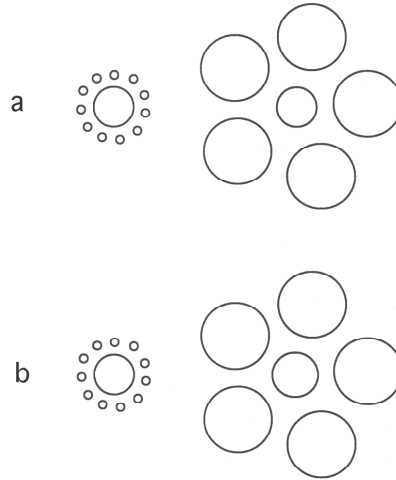


Fig. 2. Titchener (Ebbinghaus) Circles illusion. (a) The circles in the center of the two arrays perceptually appear to be different in size although they are physically identical. (b) The circles in the center of the two arrays perceptually appear to be identical in size although they are physically different.

In the experiment, Aglioti and his colleagues constructed a 3-D version of the illusion, using thin solid disks. Subjects were asked to pick up the central disk on the left if the two central disks appeared identical in size and to pick up the central disk on the right if they appeared different in size. The experimenters varied the relative size of the two target disks randomly so that in some trials physically different disks appeared perceptually identical in size, while in other trials physically identical disks appeared perceptually different in size. In selecting a disk in either trial condition, Milner and Goodale observe, “subjects indicated their susceptibility to the visual illusion” (1995/2006, 169). Nonetheless, the effect of the illusion was found to be significantly more pronounced with respect to perception (as measured by the distance between thumb and forefinger in manual estimate of disk size) than with respect to action (as measured by grip aperture in prehension). Similar findings have been reported for a variety of other visual illusions including the Müller-Lyer illusion (Daprati & Gentilucci 1997), the Ponzo illusion (Jackson & Shaw 2000), the Dot-in-Frame illusion (Bridgeman *et al.* 1997), and, recently, the Hollow-Face illusion (Króliczaka *et al.* 2006).

Milner and Goodale argue that the experimental findings provide support for the view that conscious seeing utilizes an *object-relative* metric in an allocentric or scene-based frame of reference, while visuomotor systems utilize an *absolute* metric in an egocentric frame of reference. This would explain why a pictorial, size-contrast illusion may sometimes fool the eye, but not the hand. (See Jacob & Jeannerod 2003 for a similar assessment.)

Clearly, this interpretation of the experimental findings is incompatible with EC. EC can accommodate evidence that, in addition to egocentric spatial information, conscious seeing *also* includes object- or scene-relative spatial

information (spatial information that is either not normally accessed by or less heavily weighted by visuomotor action), but it cannot accommodate evidence that conscious seeing simply does not represent the layout of visible objects and surfaces in an egocentric frame of reference.

Fortunately for EC, this interpretation is open to challenge. Although a final verdict on the comparative effects of visual illusions on action and perception is not yet in the offing, pertinent empirical considerations (to be further discussed below) include the following:

- 6.1. Many of the studies that have been cited as evidence for TV in fact indicate a theoretically significant – though comparatively less pronounced – effect of visual illusions on prehensile action. The original study by Aglioti *et al.* 1995, for instance, found that the Titchener circles illusion had a 2.5 mm effect on perception and a 1.6 mm effect on action (as measured, respectively, by the opening between index finger and thumb in a manual estimate of disk size and grip aperture). Similar findings concerning the effects of illusions on grasp position lead Ellis *et al.* 1999 to conclude that, in general, “the motor system has access to *both* the illusory perceptual information (presumably obtained from the ventral stream) and the veridical information (presumably obtained from the dorsal stream)” (1999, 113). For an overview of the relative effects of various visual illusions on action and perception, see Glover 2004.
- 6.2. In certain contexts, object-directed actions are robustly influenced by visual illusions. First, under *monocular viewing conditions*, target-directed grasping is fully affected by the Titchener Circles illusion (Goodale & Milner 2004, 92). The presumption here is that, in the absence of binocular depth information provided by stereopsis and convergence, the dorsal stream automatically “taps” pictorial depth information available in the ventral stream (Marotta *et al.* 1997, Marotta *et al.* 1998, Marotta & Goodale 1998). Second, when a *brief delay* is imposed between the disappearance of a visual target and the initiation of action in tasks involving “pantomimed” grasping or pointing, visuomotor mechanisms become fully susceptible to illusion (Goodale *et al.* 1994, Bridgeman *et al.* 1997, Hu & Goodale 2000, Westwood & Goodale 2003). The presumption here is that the longer time interval permits the dorsal stream to access spatial information temporarily stored in the ventral stream. Indeed, for this reason, visuomotor performance in optic ataxics and subjects with other forms of dorsal stream damage markedly improves with such delay (Milner *et al.* 2003, Goodale *et al.* 2004). Third, action is also highly susceptible to perceptual influence when movements are *awkward* and/or *unpracticed* (Gonzalez *et al.* 2006) and, notably, when movements are *slow* (Carey 2001, Rossetti *et al.* 2005, Króliczaka *et al.* 2006). Króliczaka *et al.* 2006 found that even the high-level Hollow-Face illusion, in which a realistic, concave mask appears to be convex when illuminated from below,¹⁶ has a strong effect on slow flicking movements directed at magnets affixed on the facing surface of the mask. Finally, Gonzalez *et al.* 2006 report that the effects of visual size illusions on grip aperture strongly depend on *which hand* is used. Grasping with the left hand was found to be fully influenced by the Titchener Circles and Ponzo illusions in both right-handed and left-handed

¹⁶ This seems to be a top-down effect in which implicit knowledge of faces and illumination conditions overrides the correct perception of concavity (indicated by stereopsis and other cues) in favor of illusory convexity and reversed 3-D depth. See Gregory 1997.

subjects. This finding suggests that the dorsal stream in the *right hemisphere* (i.e., the hemisphere contralateral to and controlling the left hand) may utilize the same sources of visuospatial information as are present in the ventral stream.

- 6.3. The studies reviewed in 6.1 and 6.2 provide evidence for the context- and task-sensitive influence of visual illusions on reaching and grasping movements. There is a significant amount of evidence however that visual illusions also have a strong influence on the programming of *saccadic eye movements*. Saccades, e.g., consistently overshoot their targets when made between the endpoints of the subjectively longer, inward-pointing segment of the Müller-Lyer illusion and consistently undershoot their targets when made between the endpoints of the subjectively shorter, outward-pointing segment (Binsted & Elliott 1999, Binsted *et al.* 2001). DiGirolamo *et al.* 2001 and McCarley & DiGirolamo 2002 have suggested that the degree of influence of the illusion on oculomotor control is based in part on the type of saccade performed. Voluntary, endogenously driven saccades are influenced by the illusion to the same degree as conscious perception. Reflexive, exogenously driven saccades, by contrast, are also influenced by the illusion, but less pronouncedly so. That said, the finding that even automatic, reflexive saccades are somewhat sensitive to pictorial visual illusions provides strong evidence for early interaction or “crosstalk” between the two putative processing streams in oculomotor control – arguably a central component of all complex visuomotor performances (returning a fast tennis serve, driving a car, running down a trail, and so on). Indeed, independent evidence for this conclusion is provided by a large body of experimental work on overt visual attention, i.e., *gaze*. A multitude of studies have found that high-level, semantic knowledge has a robust influence on the deployment of overt visual attention both when viewing a scene (Hoffman & Subramanian 1995, Rock & Mack 1998, Findlay & Gilchrist 2003) and when engaging in specific visuomotor tasks (Ballard *et al.* 1995, Hayhoe 2000, Hayhoe & Ballard 2005).¹⁷
- 6.4. Many of the contextual, pictorial depth cues that give rise to visual illusions under contrived, ecologically aberrant viewing conditions actually enhance control and guidance of visuomotor action under ecologically normal conditions, i.e., the sorts of terrestrial viewing conditions in which the human visual system evolved. Thus numerous studies have found that object-directed movements are much more accurate when made in a visually structured environment, e.g., against a textured background, than when made in a visually unstructured environment (Proteau & Masson 1997, Coello & Magne 2000, Coello & Rossetti 2004). Indeed, were the general tendency of contextual depth cues processed in the ventral stream to override or distort accurate sources 3-D spatial information independently available to visuomotor action, the evolutionary propagation of mechanisms devoted to their uptake in vision would make little biological sense. For a

¹⁷ Hubert Dreyfus has argued that “a phenomenology of expert coping shows concepts to be absent or even to get in the way of a masterful response to the specific situation” (2005, ms. 15). Studies of high-level cognitive influences on oculomotor control (in tandem with the centrality of oculomotor control to the performance of most complex visuomotor tasks), however, seem to provide compelling evidence that learned, conceptual abilities are often imbricated in the expert, non-reflective exercise of “embodied coping skills.”

brief review of studies of the role played by contextual depth cues in visuomotor action, see Dassonville & Bala 2004.

Let us now take stock.¹⁸ On the one hand, the studies reviewed in 6.1–6.4 above clearly seem to indicate that the dorsal stream (especially in the right hemisphere) has ready *access* to sources of spatial content in the ventral stream. Whether and the extent to which the dorsal stream makes use of contextual depth cues and other sources of 3-D spatial information in the ventral stream appears to vary with its task-specific needs and resources (see Clark 2001, 507-508). Indeed, points made in 6.4 suggest that accessing or “tapping” such information in the ventral stream, when feasible, generally serves to enhance control and guidance of visuomotor action. But, if this is the case, then the dorsal stream is likely to make use of spatial information in the ventral stream whenever it can afford to do so. The findings reviewed in 6.1–6.4, in short, militate against a robust, i.e., context- and task-invariant, dissociation of the two putative processing streams at the level of spatial content. They suggest a much more complicated picture, one in which the degree of interaction between the two streams depends *inter alia* on which side of the body (and, so, which hemisphere, right or left) is involved in the action, the availability to dorsal processing of its own bottom-up sources of binocular visual information, and, crucially, time constraints on performance.¹⁹

On the other hand, many of the studies reviewed in 6.1-6.3 do seem to indicate that, under certain conditions (e.g., when engaging in rapid or automatic reaching with the right hand under binocular viewing conditions or when making reflexive eye movements), visual illusions may have a measurably more *pronounced* effect on perception than on action. Why is this the case? One answer, of course, is provided by TV, by the hypothesis that visually guided action uses an egocentric frame of reference incorporating absolute metrical information, while conscious visual awareness uses an allocentric frame of reference incorporating object- or scene-relative metrical information. This hypothesis would explain why perception is less refractory than action to visual illusions involving contextual, pictorial depth cues (see Goodale & Milner 2004, 73-76). I think that a strong case for EC will have been made, if I can show that there is another hypothesis, i.e., another plausible interpretation of the studies reviewed in 6.1-6.3, that is not only consistent with EC, but that also raises fewer serious theoretical difficulties than TV. This is my objective in the following section.

¹⁸ I should note that for purposes of argument I am passing over numerous studies that strongly suggest that visual illusions actually have *identical effects* on action and perception. For relevant studies involving the Titchener Circles illusion, see Pavani *et al.* 1999, Franz *et al.* 2000, Franz 2001, Franz 2003, Franz *et al.* 2003, Vishton & Fabre 2003, and Vishton 2004. For relevant studies involving the Dot-in-Frame illusion (also known as the “Induced Roelofs Effect”), see Dassonville *et al.* 2004 and Dassonville & Bala 2004. My decision not to discuss these studies here is partially tactical. As I observe below, I think that a strong case for EC can be made by showing that there is an EC-friendly alternative to TV that both accommodates the possibility that, under certain conditions, visual illusions may have a more pronounced effect on perception than on action and also avoids serious theoretical difficulties faced by TV.

¹⁹ I am here focusing on interactions between the two streams at the level of visuospatial content. I am not including higher-level interactions mediated by stored object knowledge in the ventral stream. (See note 11 above.) The point I wish to make is that there is evidence for substantial interaction between the two streams in normal subjects even if we bracket the role played by stored object knowledge in enabling high-level, semantically rich visuomotor engagements with the world.

7. The Integration Hypothesis

According to the alternative interpretation, there is an EC-friendly explanation of why perception, i.e., conscious visual awareness, may be less refractory to visual illusion than action. The explanation is not, as proponents of TV suggest, that perception simply is not in the business of coding egocentric spatial properties, but rather that, in coding egocentric spatial properties, e.g., the distances and orientations of visible surfaces in depth, perception sometimes integrates a wider variety of (fallible) sources of spatial information than does action. In consequence, perception sometimes runs a comparatively greater risk of falling subject to visuospatial illusions.²⁰ I shall call this the “integration hypothesis.”

The integration hypothesis is supported by abundant psychophysical evidence that perception of 3-D spatial layout involves a linear, weighted averaging of independently variable sources of depth-specific information, including binocular disparity, motion parallax, occlusion, perspective, texture gradients, shading, reflections, etc. (For an overview, see Cutting & Vishton 1995 and Bruce *et al.* 2003, chap. 7.) Less theoretically contentious than the claim that egocentric spatial properties are not represented in perception is simply the hypothesis that their representation in perception involves such a weighted averaging of depth cues and that, sometimes, especially in contrived, ecologically aberrant viewing conditions, certain contextual depth cues may erroneously override or “veto” other more reliable sources of spatial information.²¹ Since the dorsal stream does not attach much relative importance to contextual depth cues in situations in which action is fast or automatic (Dijkerman *et al.* 1996, Humphrey *et al.* 1996, Marotta *et al.* 1997, Mon-Williams *et al.* 2001) it is less likely to be misled by them in those situations when these cues are inaccurate. However, when the dorsal stream’s preferred sources of spatial information are unavailable or when it has time on its hands (pun intended), the dorsal stream will make use of outputs from ventral processing and, consequently, visuomotor action will run a correspondingly greater risk of falling subject to illusion.

To sum up: Were visual illusions in fact shown sometimes to have a more pronounced effect on perception than on action, this finding would not evidence the absence of egocentric spatial coding in conscious visual awareness, as proponents of TV maintain. Rather it would simply evidence the greater sensitivity, in certain cases, of egocentric spatial coding in conscious visual awareness to potentially erroneous sources of contextual, depth-specific information (available in the ventral stream) than egocentric spatial coding in visuomotor action.

I have shown that there is an EC-friendly interpretation of the studies reviewed in 6.1-6.3 above, i.e., the integration hypothesis. I shall now proceed to

²⁰ Susanna Siegel (2006) points out that the less “committal” are the contents of visual experience, the less misperception there is. (If properties of a particular kind *F* are not represented in visual experience, then one cannot incorrectly perceive an object as having or being a certain *F*.) Conversely, the more committal are the contents of visual experience, i.e., the more varied are the kinds of properties represented in visual experience, the more misperception there is. I am making a somewhat different point: the more varied are the fallible sources of information used by the visual system to detect properties of a particular kind *F*, the more varied are the ways in which the visual system may be occasionally misinformed about the presence of *F*s (even if using more sources of information typically serves to increase the accuracy with which the visual system detects *F*s.)

²¹ It is important to emphasize once more that in ecologically normal viewing conditions contextual depth cues typically tend to corroborate – not to distort or override – other sources of 3-D spatial information. The evolutionary propagation of mechanisms devoted to their uptake in vision would otherwise make little biological sense.

show that the interpretation provided by the integration hypothesis avoids three serious theoretical difficulties that confront TV.

First, the idea that perception and action utilize fundamentally different spatial coding systems, integral to TV, gives rise to a serious problem about how the two putative systems *communicate* with one another. As Goodale and Milner write, “the two systems are using entirely different frames of reference – speaking a different language in fact – and yet somehow the ventral stream has to tell the dorsal stream which object to act upon” (2004, 101). The problem leads them to speculate that the ventral stream engages in what might be called “backward flagging.” According to this view, higher-order areas in the ventral stream working together with other cognitive systems can use back-projections to primary visual cortex (V1), the common retinotopic source of information to both streams, in order to “flag” or “highlight” targets for the dorsal stream to engage (1995/2006, 231-234; 2004, 102). Once a target has been highlighted on the retinal map in primary visual cortex, the dorsal stream can then compute its position relative to relevant parts of the body and initiate action.

It is well known that higher-order areas can prime lower-order areas in visual processing. Shifts of visual attention, as demonstrated by neurophysiological research, e.g., can actively modulate cell response patterns in the putative visual processing hierarchy as early as primary visual cortex (Ito & Gilbert 1999). There are two serious theoretical difficulties with the backward flagging view, however. The first difficulty is that many of the studies reviewed in 6.1–6.4 above point to a significant amount of high-level crosstalk or leakage between the two streams in normal subjects. Indeed, it seems clear that much more than mere retinotopic object location may be leaked to the dorsal stream from the ventral stream. In the study reported in Króliczaka *et al.* 2006, e.g., high-level knowledge of faces and normal illumination conditions stored in the ventral stream appears completely to override low-level depth information provided by binocular disparity in the dorsal stream when action is slow. Moreover, there is evidence, as we saw, that the dorsal stream in the right hemisphere (controlling action on the left side of the body) may utilize substantially the same sources of 3-D visuospatial information as are present in the ventral stream (Gonzalez *et al.* 2006). These considerations suggest that not only are there direct, high-level communication links between the two streams, but that some representational contents in the ventral stream are already in a format that the dorsal stream is able to understand.

The second theoretical difficulty is that experimental data on delayed or “pantomimed” pointing and grasping (much of it garnered by Milner and Goodale themselves) indicate that the dorsal stream is able to tap briefly stored visual representations in the ventral stream shortly after the visual target has *disappeared*. Since, in relevant cases, the object is no longer seen, there is no area on the retinotopic map in primary visual cortex corresponding to the object for the ventral stream to flag. In order to initiate action in respect of a target after its disappearance, it again seems that the dorsal stream must be able to make direct use of spatial information (stored in visual memory) in the ventral stream.

One important merit of the EC-friendly interpretation of the evidence provided by the integration hypothesis is that it bypasses the communication problem and the need to postulate something like backward flagging. Since, according to this interpretation, *both* perception, i.e., conscious visual awareness, and action make use of an egocentric frame of reference, perception has no problem when it comes to telling action upon which object to act. The testimony of the senses is delivered in a language that the body understands.

Perhaps a more significant merit of the EC-friendly interpretation is that it comports the widely accepted view that “mid-level” representation of 3-D surface

layout in a viewer-centered, i.e., egocentric, frame of reference plays a crucial role in a variety of putative *ventral* processing tasks (Marr 1982, Nakayama *et al.* 1989, Nakayama *et al.* 1995, Fleming & Anderson 2004).²² In particular, a wide array of experimental and phenomenological evidence suggests that high-level object recognition, the ventral stream's putative *raison d'être*, is significantly dependent for input on a more general purpose competence to perceive scenes in terms of surfaces egocentrically arrayed in depth. Thus, Nakayama *et al.* 1995 write: "we cannot think of object recognition as proceeding from image properties... there needs to be an explicit parsing of the image into surfaces. Without such parsing of surfaces, object recognition cannot occur" (15). If this is correct, then a third serious theoretical difficulty faced by TV is that the ventral stream in order to perform its reputed functional role must, contrary to TV, generate or have access to representations of visible surface layout in an egocentric frame of reference.

In closing this section, I should like to note that the integration hypothesis is not the only EC-friendly interpretation of the experimental evidence concerning the comparative effects of illusion on action and perception. Jeroen Smeets and Eli Brenner, in an influential series of papers, have argued that, if a single visual experience may sometimes present *inconsistent contents*, then there need not be any perceptual misrepresentation of the actual size of the central disks in the seminal Aglioti *et al.* 1995 study. In particular, the actual size of each disk may be correctly represented in visual experience even though their relative size is misrepresented as either different (figure 2a) or the same (figure 2b). Smeets and Brenner call this interpretation "the inconsistent attributes hypothesis."

As a putative example of such inconsistency Smeets and Brenner advert to the Brentano version of the Müller-Lyer illusion (figure 3). In the figure, the central vertical line *e* is correctly perceived to align with the apex of the central arrow *b* and to bisect the top horizontal line in two *equal* line segments ($ab = de = ef = bc$). At the same time, the apex of the central arrow *b* is incorrectly perceived to bisect the top horizontal line in two *unequal* parts ($ab \neq bc$).

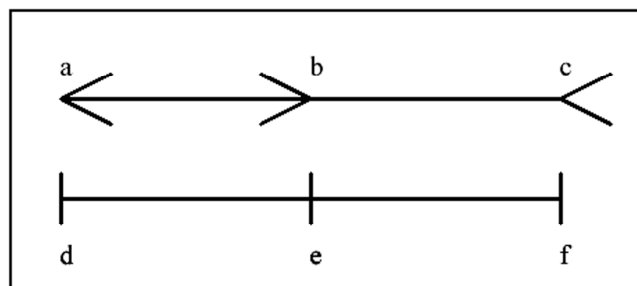


Fig. 3. Inconsistent perceptual contents illustrated by the Brentano version of the Müller-Lyer illusion. (Adapted with permission from Smeets & Brenner 2001.)

The point of the inconsistent attributes hypothesis, as Smeets and Brenner write, is that:

²² Mid-level vision is so called because it is poised in the putative visual-processing hierarchy between bottom-up, "low-level" image analysis and top-down, "high-level" object recognition. Mid-level vision represents the orientation and relative distances of non-occluded object surfaces from the observer's point of view, what is actually visible in a scene. It does not concern itself with object identities. Hence, there is an affinity between mid-level vision and what David Marr (1982) called the "2½-D sketch" of a scene.

If one realizes that various attributes of space are not necessarily represented in a consistent way, one can interpret many other experiments on illusions without the need to assume that perception and action are differentially susceptible to visual illusions. Whether an illusion affects (aspects of) the execution of a task does not depend on whether the task is perceptual or motor, but on which spatial attributes are used in (those aspects of) the task (2001, 287).

On this interpretation, the *prima facie* plausibility of which, I might note, is conceded by Clark 2001, there would be no genuine conflict between what subjects consciously *see* and what they *do* in the experiment performed by Aglioti *et al.* 1995.²³ Hence, on this interpretation, the case would not provide reason to repudiate the idea – integral to EC – that visual awareness and visually based action both utilize an egocentric spatial content base. (For further evidence in support of this interpretation and additional examples of inconsistent contents in visual experience, see de Grave *et al.* 2002, Smeets *et al.* 2002, and de Grave *et al.* 2006.)

Clark, however, in conceding the availability of the inconsistent attributes hypothesis, poses yet a third challenge to EC. “Conscious visual experience,” he writes, “may indeed present multiple inconsistent contents. But in so doing, it need not present any of those contents in a *computational format* apt for use in the fine control of online, skilled motor action” (2001, 508, my emphasis). Is there reason, however, to suppose that the computational demands made by action on the representation of spatial properties are in some sense incommensurate with how spatial properties are actually represented in visual experience? Does action require a different computational format than is used in visual experience? I address these questions and the third main source of evidence for TV in the next section.

8. The Computational Demands of Action and Perception

The proprietary biological purpose of the putative perception system, according to TV, is to create representations of objects that can be used in higher-order reasoning and in selecting goals and types of actions to perform on visual targets. The biological purpose of the putative action system, by contrast, is to enable the subject to move through the world and to engage visual targets successfully when she decides to act. “These two broad objectives,” Goodale and Milner write, “impose such conflicting requirements on the brain that to deal with them within a single unitary visual system would present a computational nightmare” (2004, 73).

These considerations, in tandem with other the empirical evidence reviewed above, suggest to proponents of TV that in visual awareness we are delivered with only coarse-grained, relative metrical information about objects in an allocentric or scene-based frame reference. By contrast, visuomotor control relies on fine-grained,

²³ The positing of visual experiences with inconsistent contents is not merely an *ad hoc* move to safeguard EC. It can be independently motivated, as suggested by figure 3. The well-known “waterfall illusion,” I might note, also seems to involve inconsistent contents in visual experience. (See Crane 1988 for an argument to this effect.) After fixating for a length of time on the downward flow in a waterfall, one may have the strong visual impression, induced by perceptual adaptation to the movement, that stationary objects, e.g., trees on the riverbank, are moving upwards. Objects appear both to be stationary *and* to move at the same time. Thanks to an anonymous referee for raising concerns that required me to clarify this point.

absolute metrical information about objects in an egocentric frame of reference. Thus Clark writes:

...fine-grained action control requires the extraction and use of *radically different kinds of information* (from the incoming visual signal) than does recognition, recall, and reasoning. The former requires a constantly updated, egocentrically-specified, exquisitely distance- and orientation-sensitive encoding of the visual array. The latter requires a degree of object-constancy, and the recognition of items by category and significance irrespective of the fine detail of location, viewpoint and retinal image size. A computationally efficient coding for either task precludes the use of the very same encoding for the other... (2001, 501, my emphasis).

My first observation in this connection is that the question as to whether fine-grained visuomotor control requires “radically different kinds of information” than is required for purposes of *recognition, recall, and reasoning* is surely distinct from the question as to whether visuomotor control requires radically different spatial information than is made available to us in visual awareness. It may well be the case that different information is required for purposes, say, of putting a kettle on the stove than is required for purposes of kettle identification or for purposes of thinking kettle-related thoughts. But this by itself would not warrant the conclusion that the *spatial properties* of the kettle (its location, orientation, distance, etc.) are represented in a significantly different way by putative systems subserving visual awareness than by putative systems subserving visuomotor action.

My second observation is that clearly we do not perceive only the sizes and positions of objects relative to one another in a scene. We also accurately and reliably perceive their intrinsic spatial properties and, in particular, their spatial relations to *us*, e.g., their respective egocentric directions and distances in depth. This, I take it, is an uncontroversial phenomenological claim about the deliverances of normal, spatially contentful perceptual experience. But there is a stronger epistemological claim in the offing. This would be the claim is that it is actually a *necessary condition* of the possibility of spatially contentful perceptual experience in general that one have abilities to perceive the spatial relations of objects to one’s own bodily location. What would it be, one might ask, to have the abilities needed perceptually to identify the locations of objects relative to other objects in one’s field of view – e.g., the location of one candlestick relative to another on the dining table – but *not* those needed ever to identify the location of objects relative to one’s own body? Is such non-perspectival, i.e., purely allocentric, experience of space possible? Can we coherently imagine a being that possessed the former abilities, but not the latter? If not, is the reason that perceptual abilities for allocentric identification of location are actually dependent on perceptual abilities for egocentric identification of location?²⁴

(An object’s location may be given indirectly, of course, through an appropriate, demonstratively anchored description (Brewer 1997, 190). In reporting one’s visual experience, one may describe a candlestick as *on top of that table*, or *next to that saltcellar*, etc. These, however, do not seem to be pure cases of allocentric identification since they plausibly depend on direct perception of an object’s

²⁴ Plausibly a main upshot of the Kantian-flavored argument in the first part of Strawson 1959 is that no identification of location in an objective spatio-temporal order would be possible were it not for more basic perceptual capacities to identify the locations of objects in a subject-centered frame of reference. For elaboration of Strawson’s thesis, see Evans 1982, chap. 6 and Grush 2000.

location relative to one's own. The candlestick, e.g., is perceived as next to *that* saltcellar, where *that* saltcellar is located in *front* of one, *down* a little, and over to the *left*.)

Goodale and Milner make much of the fact that we are easily able to perceive the relative locations of objects depicted in films and photographs (2004, 73-76). We easily see, for instance, that Asta is peeking out from *under* the bed *beside* the bureau *in front* of which William Powell is standing in a scene from *The Thin Man* despite the fact that we do not stand in any actual egocentric spatial relations to the objects that figure in our perception. This sort of observation, in conjunction with findings about the influence of size-contrast illusions in grasping tasks, suggests to Goodale and Milner that visual awareness exploits an essentially scene-relative frame of reference. I would suggest, however, that such perceptual identification of object locations in films and photographs crucially involves a kind of hypothetical perceptual identification of location relative to one's own body inasmuch as, when watching a film or looking at a photograph, one non-reflectively assumes the perspective of the camera.²⁵

As Ruth Millikan points out, "That human perceivers can retrieve information from photographs and television depends on their capacity to use information about distal affairs that are not represented or yet understood as having definite and useful relations *to themselves*" (2004, 123). I am suggesting here, however, that having this capacity depends on being able implicitly to assume a certain hypothetical point of view in respect of the imaged scene. In viewing a portrait, I may not know where on earth the person portrayed sat in order to be photographed, but I do know that, had *I* taken the picture, I would have been facing the person and standing slightly to her right. Spatially contentful visual experience, I would suggest, is always perspectival in this way (see Brewer 1997, chap. 6).

One main reason Goodale and Humphrey give for the claim that the informational demands of visual awareness and visuomotor control are dramatically different is that were perceptual representations underlying visual awareness "to attempt to deliver the real metrics of all objects in the visual array, the computational load would be astronomical" (1998, 195-96). Similarly, Goodale and Milner write, "To have to take in the absolute metrics of the entire scene would in fact be computationally impossible, given the rapidity with which the pattern of light changes on our retina" (2004, 82). Such observations lead Clark to conclude that "Visual awareness, if this story is correct, cannot *afford* to be action oriented... It simply cannot afford to represent each and every aspect of the scene present in visual awareness in the precise and egocentrically defined co-ordinates required to support complex physical interactions with that very scene" (1999).

But why ought we to assume that, were visual awareness to represent objects using real-world metrics in an egocentric frame of reference, then it would have to do so simultaneously for "*all* objects in the visually array," for "*the entire scene*," and for "*each and every aspect*" in it? It seems clear that visuomotor control must make highly focused, task-specific use of metrically precise visual information (Ballard *et al.* 1997, Findlay & Gilchrist 2003). It certainly does not require a metrically precise model of the *entire* scene. Why, then, should not the same hold true for visual awareness as well? Indeed, as Goodale and Milner (2004, 94-96) observe, the widely studied phenomenon of *change blindness* suggests that visual awareness of a complete, coherent, and uniformly detailed world does not require access to a complete, coherent, and uniformly detailed internal representation of the

²⁵ In his commentary on *The Varieties of Reference*, Rick Grush refers to this as "mock egocentric" identification (mind.ucsd.edu/misc/resources/evans/evansindex.html).

world.²⁶ Change blindness suggests that which details we notice in a scene – and, so, which details may vary without notice – is a function of visual attention. At any given moment, very little information about the scene may be retained in short-term visual memory. Representation of rich detail, however, doesn't require access to richly detailed representations in the mind. The detail is present in the world itself and, so, the world can serve as "its own best model" (Brooks 1991) or as an "outside memory" (O'Regan 1992).²⁷ In short, EC can avail itself of the intuitively tenable assumption that in visual awareness we are delivered with precise – if local and attentionally constrained – metrical information about objects in a perspectival, egocentric frame of reference while forgoing the intuitively untenable assumption that at any given moment we are delivered with precise information about the real metrics of each and every object in the scene.

It is important to make two points in concluding this section. First, the key claim that personal-level visual awareness and visually based action both utilize an egocentric spatial content base depends for its plausibility on clearly distinguishing between representational *content* and representational *vehicle* (Dennett 1991, Millikan 1993). What is common to visual awareness and visually based action, according to EC, is to be typed at the level of *content*. Hence, conscious seeing can deliver us with the same kind of egocentric spatial content as is utilized in visuomotor action, even if the representational *vehicles* that respectively support conscious seeing and visuomotor action are sometimes physically different. It should be noted in this connection that, in defending EC, it is possible to remain noncommittal about the relative contributions of subpersonal ventral and dorsal stream processing to personal-level visuospatial awareness in healthy subjects. In particular, it is possible to abstain from Milner and Goodale's theoretical characterization of the ventral stream as a functionally independent "vision for perception" system. While it seems clear that dorsal stream processing by itself is functionally insufficient for normal visuospatial awareness (as evidenced in part by the profound visual form agnosia consequent upon trauma to the ventral stream in subject DF), there is significant neuropsychological evidence – briefly discussed at the end of §5 – that much dorsal stream processing may nonetheless be functionally necessary. The claim that visual awareness makes use of an egocentric spatial content base, accordingly, should not be understood as the claim that the ventral stream considered in *isolation*, i.e., apart from its interactions with the dorsal stream and higher-order, "executive" areas in the brain, causally supports or enables visual awareness of egocentric space.

"The representational content of experience," Peacocke writes, "is a many-splendored thing. This is true not only of the remarkable range of detail in the perceptual content but also of the range of different and philosophically interesting types of content that can be possessed by a particular experience" (1992, 61). The second point is that, in keeping with this observation, EC does not restrict the spatial representational contents of visual awareness to egocentric spatial properties. It does not deny that visual awareness may *also* convey information to the subject about a wide assortment of allocentric or scene-relative spatial properties. When I view two candlesticks on the dining table, e.g., I see not only their relative distances in depth from me, but also their spatial relations to one

²⁶ See Rensink 2000, Noë 2002, Rensink & Simons 2005.

²⁷ To make this observation, however, is not to endorse the view that *all* representation of detail is "virtual" as Alva Noë (2004, 134) proposes. Which details succeed in capturing attention, including which changes are noticed, for instance, crucially depends on the broad meaning or "gist" of the scene, and visual recognition of gist depends, in turn, on rapid processing and integration of represented spatial details across the scene. See Mack & Rock 1998; Mack 2002; and Koch 2004, chap. 9.

another and to other visible objects and surfaces in the room. Notably, certain objects have intrinsic axes of their own in relation to which I am able to perceive the locations of things around them. When I view a cat, e.g., I see not only its location relative to my bodily axes, but also the locations of other objects to *its* intrinsic axes. Whether a toy mouse is perceived as to the right, or in front of, or behind the cat depends on the cat's orientation relative to the toy, not my own. EC claims only that visual experience represents 3-D space using an egocentric frame of reference, i.e., using multiple, coordinated effector-specific frames of reference. It does not deny that visual experience may also utilize object- and scene-relative frames of reference.

9. Perceptual Consciousness and Rational Agency

In closing, I would like to make two related observations. According to EC, a subject is perceptually aware of the region occupied by an object in egocentric space to the extent that she has a practical understanding of the way her bodily spatial relations to the object would vary with possible intentional movements and actions (§2). The first observation is that EC is thus consistent with the view shared by many researchers that there is a profound connection between perceptual consciousness and rational agency (Marcel 1988, Humphrey 1992, McDowell 1994, Brewer 1997, Gallagher 2005, Dretske 2006). In particular, EC is consistent with the view that visually transduced spatial information is conscious only when it is available to the subject as a potential reason for voluntary, intentional action. As Fred Dretske writes, "It is the availability of [such] information for rationalizing and motivating intentional action (even if one is not capable of such action – e.g., paralyzed or buried in cement), not its actual use, that makes it conscious" (2006, 174). What gives sense to the claim that in consciously seeing, say, that an object *o* is situated in a certain egocentrically identified location *l* one is delivered with a *potential reason* for action is simply the fact that, were it one's intention at the time of the visual experience to move toward *o*, or to turn toward *o*, or to point toward *o*, etc., then *l* would be the location, *ceteris paribus*, toward which it would be rational to move, or to turn, or to point. A description of the spatial information conveyed by one's visual experience would be an essential part of a complete, intentional explanation of the action one would be motivated to perform.²⁸ According to EC, visually transduced egocentric spatial information is conscious – available to the subject herself and not just to her visual system – to the extent that it is able rationally to inform the subject's intentional, bodily actions in this way.²⁹

The second observation is that it is important for this view that it be able to show that the determinacy of detail or "fineness of grain" with which egocentric spatial properties are represented in visual experience does not outstrip the

²⁸ The case, I take it, is comparable to the case in which one's action is justified in part by a *belief* about an object's location. What makes going to the kitchen for some milk fully rational is not only one's desire for milk, but also one's belief that the kitchen is *where* milk is to be found.

²⁹ Two points: First, this is not to deny that such information will also be available for purposes of reasoning, imagining, and so on. Thanks to Ruth Millikan for emphasizing the need to make this point. Second, formulating the claim in this way raises the question whether a sharp theoretical distinction can be drawn between intentional action and intelligent, environmentally adaptive behavior; between manifestations of the subject's own purposes and manifestations of the (biological) purposes of her subpersonal parts. Like others, I am skeptical that a general theoretical distinction that is both sharp and useful can be drawn here (Dennett 1991; Gazzaniga 1998; Hurley 1998; Koch 2004; Millikan 2004, chap. 1).

determinacy of detail with which they are potentially represented in the contents of one's intentions for spatially directed movement and action. Suppose that a phone is ringing in a busy hotel lobby, and, on a whim, one stops in passing to answer it. It seems fairly clear that one might vaguely notice the location of the receiver, with the rest happening more or less automatically. In connection with this no doubt ubiquitous sort of case, Campbell's analogy with a heat-seeking missile (§4) is fitting. In such a case, we may suppose that once the visuomotor system has "locked" onto the gross location of the phone, metrically more precise visual information about its shape, size, and orientation is then nonconsciously used to determine the trajectory taken by one's arm and the scaling of one's hand to the receiver.

Obviously, however, not all of our intentions to engage with visual targets are based on a mere awareness of their general location. Consider, in contrast, the case in which a jeweler delicately etches the crossbar of the letter "t" in an inscription on wedding ring; or the case in which a parent gently removes a tiny splinter from a child's finger; or the case in which one visually tracks a floating mote of dust with the movements of one's eyes. These, to be sure, are all cases of unusually fine, intentional visuomotor control. But they illustrate the general point, I take it, that *any* of the spatial details one visually experiences in a scene may provide one with a reason for action. (Whether a visually experienced detail provides one with an actual reason for action, of course, will depend on one's current motivations, aims, beliefs, etc.) If this is the case, however, then there is no need to worry that the determinacy of detail with which spatial properties are represented in visual experience may outstrip the determinacy of detail with which they are potentially represented in the contents of one's intentions for spatially directed movement and action. The spatial resolution of the contents of those intentions can be just as sharp as the spatial resolution of vision itself.

References

- Aglioti, S., M. Goodale, & J. DeSouza. 1995. "Size Contrast Illusions Deceive the Eye But Not the Hand." *Current Biology* 5: 679-85.
- Ballard, D., M. Hayhoe, & J. Pelz. 1995. "Memory Representations in Natural Tasks." *Journal of Cognitive Neuroscience* 7: 66-80.
- Ballard, D., M. Hayhoe, P. Pook, & R. Rao. 1997. "Deictic Codes for the Embodiment of Cognition." *Behavioral and Brain Sciences* 20: 723-67.
- Battaglia-Mayer, A. & R. Caminiti. 2002. "Optic Ataxia as the Result of the Breakdown of the Global Tuning Fields of Parietal Neurones." *Brain* 125: 225-237.
- Bermúdez, J. 1998. *The Paradox of Self-Consciousness*. Cambridge, MA: MIT Press.
- Binsted, G. & D. Elliott. 1999. "The Müller-Lyer Illusion as a Perturbation to the Saccadic System." *Human Movement Science* 18: 103-17.
- Binsted, G., D. Elliott, W. Helsen, & R. Chua. 2001. "Eye-hand Coordination in Goal-directed Aiming." *Human Movement Science* 20: 563-85.
- Blakemore, S., D. Wolpert, & C. Frith. 2002. "Abnormalities in the Awareness of Action." *Trends in Cognitive Sciences* 6: 237-242.
- Blangero, A., H. Ota, L. Delporte, P. Revol, P. Vindras, G. Rode, D. Boisson, A. Vighetto, Y.

- Rossetti, & L. Pisella. 2007. "Optic Ataxia is Not Only 'Optic': Impaired Spatial Integration of Proprioceptive Information." *Neuroimage* 36: T61–T68.
- Brenner, E. & J. Smeets. 2001. "Action Beyond Our Grasp." *Trends in Cognitive Sciences* 5: 287.
- Brewer, B. 1995. "Bodily Awareness and the Self." In J. Bermúdez, A. Marcel, and N. Eilan, eds., *The Body and the Self*. Cambridge, MA: MIT Press.
- 1997. *Perception and Reason*. Oxford: Oxford University Press.
- Brooks, R. 1991. "Intelligence Without Representation." *Artificial Intelligence* 47: 139–159.
- Bruce, V., P. Greene, & M. Georgeson, eds. 2003. *Visual Perception: Physiology, Psychology, and Ecology*, 4th ed. London: Psychology Press.
- Campbell, J. 1994. *Past, Space, and Self*. Cambridge, MA: MIT Press.
- 2002. *Reference and Consciousness*. Oxford: Oxford University Press.
- Carey, D. 2001. "Do Action Systems Resist Visual Illusions?" *Trends in Cognitive Sciences* 5: 109–113.
- Clark, A. 1999. "Visual Awareness and Visuomotor Action." *Journal of Consciousness Studies* 6: 1–18.
- 2001. "Visual Experience and Motor Action: Are the Bonds Too Tight?" *The Philosophical Review* 110: 495–519.
- Coello, Y. & P. Magne. 2000. "Determination of Target Position in a Structured Environment: Selection of Information for Action." *European Journal of Cognitive Psychology* 12: 489–519.
- Coello, Y. & Y. Rossetti. 2000. "Planning and Controlling Action in a Structured Environment: Visual Illusion Without Dorsal Stream." *Behavioral and Brain Sciences* 27: 29–31.
- Cohen, Y. & R. Andersen. 2002. "A Common Reference Frame for Movement Plans in the Posterior Parietal Cortex." *Nature Neuroscience* 3: 553–562.
- Colby, C. 1998. "Action-oriented Spatial Reference Frames in Cortex." *Neuron* 20: 15–24.
- Colby, C. & M. Goldberg. 1999. "Space and Attention in Parietal Cortex." *Annual Review of Neuroscience* 22: 319–49.
- Crane, T. 1988. "The Waterfall Illusion." *Analysis* 48: 142–147.
- Cussins, A. 2003. "Experience, Thought, and Activity." In Y. Gunther, ed., *Essays on Nonconceptual Content*. Cambridge, MA: MIT Press.
- Cutting, J. & P. Vishton. 1995. "Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth." In W. Epstein & S. Rogers, eds., *Perception of Space and Motion*. San Diego: Academic Press.
- Daprati, E. & M. Gentilucci. 1997. "Grasping an illusion." *Neuropsychologia* 35: 1577–1582.
- Dassonville, P. & J. Bala. 2004. "Perception, Action, and Roelofs Effect: A Mere Illusion of Dissociation." *PLoS Biology* 2:11: e364: 1936–1945.
- Dassonville, P., B. Bridgeman, J. Bala, P. Thiem, & A. Sampanes. 2004. "The Induced Roelofs Effect: Two Visual systems or the Shift of a Single Reference Frame?" *Vision Research* 44: 603–

611.

de Grave, D., J. Smeets, & E. Brenner. 2002. "Ecological and Constructivist Approaches and the Influence of Illusions." *Behavioral and Brain Sciences* 25: 103-104.

—2006. "Why Are Saccades Influenced by the Brentano Illusion?" *Experimental Brain Research* 175: 177–182.

Dennett, D. 1991. *Consciousness Explained*. Boston: Little, Brown, and Co.

Dijkerman, H., A. Milner, & D. Carey. 1996. "The Perception and Prehension of Objects Oriented in the Depth Plane. I. Effects of Visual Form Agnosia." *Experimental Brain Research* 112: 442–51.

Dretske, F. 2006. "Perception Without Awareness." In Gendler & Hawthorne 2006.

Dreyfus, H. 2005. "Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise." APA Pacific Division Presidential Address.

Eilan, N., R. McCarthy, & B. Brewer, eds. 1993. *Spatial Representation*. Oxford: Oxford University Press.

Ellis, R., J. Flanagan, & S. Lederman. 1999. "The Influence of Visual Illusions on Grasp Position." *Experimental Brain Research* 125: 109-114.

Evans, G. 1982. *The Varieties of Reference*. J. McDowell, ed. Oxford: Oxford University Press.

—1985. "Molyneux's Question." In A. Phillips, ed., *The Collected Papers of Gareth Evans* London: Oxford University Press.

Findlay, J. & I. Gilchrist. 2003. *Active Vision: The Psychology of Looking and Seeing*. Oxford: Oxford University Press.

Fleming, R. & B. Anderson. 2004. "The Perceptual Organization of Depth." In L. Chalupa and J. Werner, eds., *The Visual Neurosciences*. Cambridge, MA: MIT Press.

Franz, V. 2001. "Action Does Not Resist Visual Illusions." *Trends in Cognitive Sciences* 5: 457–459.

—2003. "Manual Size Estimation: A Neuropsychological Measure of Perception?" *Experimental Brain Research* 151: 471–477.

Franz, V., K. Gegenfurtner, H. Bühlhoff, & M. Fahle. 2000. "Grasping Visual Illusions: No Evidence for a Dissociation Between Perception and Action." *Psychological Science* 11: 20–25.

Franz, V. H. Bühlhoff, & M. Fahle. 2003. "Grasp Effects of the Ebbinghaus Illusion: Obstacle Avoidance is Not the Explanation." *Experimental Brain Research* 149: 470–477.

Gallagher, S. 2005. *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gallagher, S. & A. Marcel. 1999. "The Self in Contextualized Action." *Journal of Consciousness Studies* 6: 4-30.

Gallagher, S. & M. Jeannerod. 2002. "From Action to Interaction." *Journal of Consciousness Studies* 9: 3-26.

Gallese, V. 2005. "Embodied Simulation: From Neurons to Phenomenal Experience." *Phenomenology and the Cognitive Sciences* 4: 23-48.

- 2007. "The 'Conscious' Dorsal Stream: Embodied Simulation and its Role in Space and Action Conscious Awareness." *Psyche* 13 (<http://psyche.cs.monash.edu.au/>).
- Gazzaniga, M. 1998. *The Mind's Past*. Berkeley: The University of California Press.
- Gendler, T. & J. Hawthorne, eds. 2006. *Perceptual Experience*. Oxford: Oxford University Press.
- Gibson, J. 1966. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.
- 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Glover, S. 2004. "Separate Visual Representations in the Planning and Control of Action." *Behavioral and Brain Sciences* 27: 3–78.
- Gonzalez, C., T. Ganel, & M. Goodale. 2006. "Hemispheric Specialization for the Visual Control of Action Is Independent of Handedness." *Journal Neurophysiology* 95: 3496-3501.
- Goodale, M. 1998. "Visuomotor Control: Where Does Vision End and Action Begin?" *Current Biology* 8: R489-R491.
- Goodale, M. & K. Humphrey. 1998. "The Objects of Action and Perception." *Cognition* 67: 181-207.
- Goodale, M. & D. Milner. 2004. *Sight Unseen*. Oxford: Oxford University Press.
- Goodale, M., L. Jakobson, & J. Keillor. 1994. "Differences in the Visual Control of Pantomimed and Natural Grasping Movements." *Neuropsychologia* 32: 1159–78.
- Goodale, M., D. Westwood, & A. Milner. 2004. "Two Distinct Modes of Control for Object-directed Action." *Progress in Brain Research* 144: 131–44.
- Gregory, R. 1997. "Knowledge in Perception and Illusion." *Philosophical Transactions of the Royal Society of London B* 352: 1121–1128.
- Grush, R. 1998. "Skill and Spatial Content." *Electronic Journal of Analytic Philosophy* 6.
- 2000. "Self, World, and Space: The Meaning and Mechanisms of Ego- and Allocentric Spatial Representation." *Brain and Mind* 1: 59–92.
- Hayhoe, M. 2000. "Vision Using Routines." *Visual Cognition* 7: 43-64.
- Hayhoe, M. & D. Ballard 2005. "Eye Movements in Natural Behavior." *Trends in Cognitive Sciences* 9: 188-194.
- Heck, R. *forthcoming*. "Are There Different Kinds of Content?"
- Hoffman, J. & B. Subramanian. 1995. "Saccadic Eye Movements and Visual Selective Attention." *Perception and Psychophysics* 57: 787–95.
- Hu, Y. & M. Goodale 2000. "Grasping After a Delay Shifts Size-scaling From Absolute to Relative Metrics." *Journal of Cognitive Neuroscience* 12: 856-68.
- Humphrey, N. 1992. *A History of the Mind*. New York: Simon and Schuster.
- Humphrey, G., L. Symons, A. Herbert, & M. Goodale. 1996. "A Neurological Dissociation Between Shape from Shading and Shape from Edges." *Behavioural Brain Research* 76: 117–25.

- Hurley, S. 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- Ito, M. & C. Gilbert. 1999. "Attention Modulates Contextual Influences in the Primary Visual Cortex of Alert Monkeys." *Neuron* 22: 593-604.
- Jackson, S. & A. Shaw. 2000. "The Ponzio Illusion Affects Grip-force But Not Grip-aperture Scaling During Prehension Movements." *Journal of Experimental Psychology HPP* 26: 418-23.
- Jacob, P. & M. Jeannerod. 2003. *Ways of Seeing*. Oxford: Oxford University Press.
- Jeannerod, M. 1997. *The Cognitive Neuroscience of Action*. Oxford: Blackwell.
- Kelly, S. 2004. "On Seeing Things in Merleau-Ponty." In T. Carmon, ed., *The Cambridge Companion to Merleau-Ponty*. Cambridge: Cambridge University Press.
- Knox, J., M. Coppieters, & P. Hodges. 2006. "Do You Know Where Your Arm Is if You Think Your Head Has Moved?" *Experimental Brain Research* 173: 94-101.
- Koch, C. 2004. *The Quest for Consciousness*. Englewood: Roberts and Co.
- Króliczaka, G., P. Heard, M. Goodale, & R. Gregory. 2006. "Dissociation of Perception and Action Unmasked by the Hollow-face Illusion." *Brain Research* 1080: 9-16.
- Land, M. & S. Furneaux. 1997. "The Knowledge Base of the Oculomotor System." *Philosophical Transactions of the Royal Society of London B* 352: 1231-1239.
- Land, M. & P. McLeod. 2000. "From Eye Movements to Actions: How Batsmen Hit the Ball." *Nature Neuroscience* 3: 1340-1345.
- Mack, A. 2002. "Is the Visual World a Grand Illusion? A Response." *Journal of Consciousness Studies* 9: 102-10.
- Mack, A. & I. Rock. 1998 *Inattentional Blindness*. Cambridge, MA: MIT Press.
- Mandik, P. 1999. "Qualia, Space, and Control." *Philosophical Psychology* 12:1:47-60.
- Marcel, A. 1988. "Phenomenal Experience and Functionalism." In A. Marcel and E. Bisiach, eds., *Consciousness in Modern Science*. Oxford: Oxford University Press.
- 2003. "The Sense of Agency: Awareness and Ownership of Action." In Roessler & Eilan 2003.
- Marotta, J., M. Behrmann, & M. Goodale. 1997. "The Removal of Binocular Cues Disrupts the Calibration of Grasping in Patients with Visual Form Agnosia." *Experimental Brain Research* 116: 113-21.
- Marotta, J., J. DeSouza, A. Haffenden, & M. Goodale. 1998. "Does a Monocularly Presented Size-Contrast Illusion Influence Grip Aperture?" *Neuropsychologia* 36: 491-97.
- Marotta, J. & M. Goodale. 1998. "The Role of Learned Pictorial Cues in the Programming and Control of Grasping." *Experimental Brain Research* 121: 465-70.
- Marr, D. 1982. *Vision*. New York: W.H. Freeman & Sons.
- McCarley, J. & G. DiGirolamo. 2002. "One Visual system with Two Interacting Visual Streams." *Behavioral and Brain Sciences* 24: 112-113.

- McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- Merleau-Ponty, M. 1945/1962. *The Phenomenology of Perception*. C. Smith, transl. London: Routledge and Kegan Paul Ltd.
- Millikan, R. 1993. "Content and Vehicle." In Eilan *et al.* 1993.
- 2004. *Varieties of Meaning*. Cambridge, MA: MIT Press.
- 2006. "Postulating Perceptual Representations in a Way that Actually Supports Gibson's Central Insights." Lecture at the Boston Colloquium for Philosophy of Science, January 30, 2006.
- Milner, D. & M. Goodale. 1995/2006. *The Visual Brain in Action*, 2nd edition. Oxford: Oxford University Press.
- Milner, A., H. Dijkerman, R. McIntosh, Y. Rossetti, & L. Pisella. 2003. "Delayed Reaching and Grasping in Patients with Optic Ataxia." *Progress in Brain Research* 142: 225-42.
- Mon-Williams, M., J. Tresilian, R. McIntosh, & A. Milner. 2001. "Monocular and Binocular Distance Cues: Insights from Visual Form Agnosia I (of III)." *Experimental Brain Research* 139: 127-136.
- Moreaud, O. 2003. "Balint Syndrome." *Archives of Neurology* 60: 1329-1331.
- Nakayama, K., S. Shimojo, G. Silverman. 1989. "Stereoscopic Depth: Its Relation to Image Segmentation, Grouping, and the Recognition of Occluded Objects." *Perception* 18: 55-68.
- Nakayama, K., Z. He, & S. Shimojo. 1995. "Visual Surface Representation: A Critical Link Between Lower-level and Higher-level Vision." In S. Kosslyn & D. Osherson, eds., *Visual Cognition*. Cambridge, MA: MIT Press.
- Nielsen, T. 1963. "Volition: A New Experimental Approach." *Scandinavian Journal of Psychology* 4: 225-230.
- Noë, A., ed. 2002. "Is the Visual World a Grand Illusion?" Thorverton: Imprint Academic.
- 2004. *Action in Perception*. Cambridge, MA: MIT Press.
- O'Regan, K. 1992. "Solving the 'Real' Mysteries of Visual Perception: The World as an Outside Memory." *Canadian Journal of Psychology* 46: 461-488.
- O'Shaughnessy, B. 1980. *The Will (Volume I)*. Cambridge: Cambridge University Press.
- Paillard, J. 1991. "Motor and Representational Framing of Space." In J. Paillard, ed., *Brain and Space*. Oxford: Oxford University Press.
- Palmer, S. 1999. *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.
- Pavani, F., I. Boscagli, F. Benvenuti, M. Rabuffetti, & A. Farnè. 1999. "Are Perception and Action Affected Differently by the Titchener Circles Illusion?" *Experimental Brain Research* 127: 95-101.
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge, MA: MIT Press.
- 1993. "Intuitive Mechanics, Psychological Reality and the Idea of a Material Object." In Eilan *et al.* 1993.

- Perenin, M. & A. Vighetto. 1988. "Optic Ataxia: A Specific Disruption in Visuomotor Mechanisms." *Brain* 111: 643-674.
- Pesaran, B., M. Nelson, & R. Andersen. 2006. "Dorsal Premotor Neurons Encode the Relative Position of the Hand, Eye, and Goal During Reach Planning." *Neuron* 51: 125-134.
- Prinz, W. 1990. "A Common Coding Approach to Perception and Action." In O. Neumann and W. Prinz, eds., *Relationships Between Perception and Action: New Approaches*. Berlin: Springer.
- Proteau, L. & G. Masson. 1997. "Visual Perception Modifies Goal-Directed Movement Control: Supporting Evidence from a Visual Perturbation Paradigm." *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 50A: 726-41.
- Rensink, R. 2000. "The Dynamic Representation of Scenes." *Visual Cognition* 7: 17-42.
- Rescorla, M. 2003. *Is Thought Explanatorily Prior to Language?* Doctoral Dissertation. Harvard University, Cambridge, MA.
- Rizzolatti, G., L. Fadiga, L. Fogassi, & V. Gallese. 1997. "The Space Around Us." *Science* 277: 190-191.
- Rizzolatti, G. & M. Matelli. 2003. "Two Different Streams Form the Dorsal Visual System: Anatomy and Functions." *Experimental Brain Research* 153: 146-157.
- Roessler, J. & N. Eilan, eds. 2003. *Agency and Self-Awareness: Issues in Philosophy and Psychology*. Oxford: Oxford University Press.
- Rossetti, Y., L. Pisella, & A. Vighetto. 2003. "Optic Ataxia Revisited: Visually Guided Action Versus Immediate Visuomotor control." *Experimental Brain Research* 153: 171-179.
- Rossetti, Y., P. Revol, R. McIntosh, L. Pisella, G. Rode, J. Danckert, C. Tilikete, H. Dijkerman, D. Boisson, A. Vighetto, F. Michel, & A. Milner. 2005. "Visually Guided Reaching: Bilateral Posterior Parietal Lesions Cause a Switch from Fast Visuomotor to Slow Cognitive Control." *Neuropsychologia* 43: 162-177.
- Siegel, S. 2006. "Which Properties are Represented in Perception?" In Gendler & Hawthorne 2006.
- Simons, D. & R. Rensink. 2005. "Change Blindness: Past, Present, and Future." *Trends in Cognitive Sciences* 9: 16-20.
- Smeets, J. & E. Brenner. 2001. "Action Beyond Our Grasp." *Trends in Cognitive Sciences* 5: 287.
- Smeets, J., E. Brenner, D. Grave, & R. Cuijpers. 2002. "Illusions in Action: Consequences of Inconsistent Processing of Spatial Attributes." *Experimental Brain Research* 147: 135-144.
- Sørensen, J. 2005. "The Alien-Hand Experiment." *Phenomenology and the Cognitive Sciences* 4: 73-90.
- Strawson, P. 1959. *Individuals*. London: Methuen and Co. Ltd.
- Vishton, P. 2004. "Human Vision Focuses on Information Relevant to a Task, to the Detriment of Information That is Not Relevant." *Behavioral and Brain Sciences* 27: 53-54.
- Vishton, P. & E. Farbe. 2003. "Effects of the Ebbinghaus Illusion on Different Behaviors." *Spatial Vision* 16: 377-92.

Weiskrantz, L., E. Warrington, M. Saunders, & J. Marshall. 1974. "Visual Capacity in the Hemianopic Field Following a Restricted Occipital Ablation." *Brain* 97: 709-28.

Westwood, D. & M. Goodale. 2003. "Perceptual Illusion and the Real-time Control of Action." *Spatial Vision* 16: 243-54.