

Grasping and perceiving objects¹

Pierre Jacob
Institut Jean Nicod,
CNRS/EHESS/ENS,
Paris, France

Introduction

There is today, or so I will argue in this paper, a vast array of empirical evidence in favor of the ‘two visual systems’ model of human vision. Human beings are so visually endowed that they can see a wide variety of things. Some of the things that they can see are objects that they can also reach, grasp and manipulate with their hands. Many of the things that they can see, however, are not objects that they can reach and grasp.² As Austin (1962) pointed out, humans can see e.g., mountains, lakes, liquids, gases, clouds, flames, movies, shadows, holes, stars, planets, comets, and events. Among events, humans can see behaviors or actions, some of which are performed by conspecifics. Some visible human actions are directed towards inanimate objects (e.g., actions of reaching and grasping an object or a tool). Others are directed towards animate objects, including animals and conspecifics. In Jacob and Jeannerod (in press), we argue that there is evidence for the view that the human brain

¹ First, I am grateful to Kathleen Akins, Steven Davis and especially Andrew Brook, who organized the wonderful Conference on Philosophy and the Neurosciences at Carleton University, October 17-20, 2002. Secondly, I am also grateful to Andrew Brook and especially to Kathleen Akins for penetrating comments on this chapter. Finally, I wish to register my deep gratitude to Marc Jeannerod, who introduced me to the cognitive neurosciences of vision and action.

² Nor are actions that humans can perform reducible to manual actions of prehension.

contains two complementary networks that respond to the perception of respectively object-oriented actions and actions directed towards conspecifics.

In this paper, I will restrict myself to the visual processing of objects that can be both perceived and grasped with a human hand. My goal will be to try to formulate an adequate version of the two visual systems model of human vision, which ought, I think, to be properly restricted to seeing objects that can be reached and grasped. At the heart of my version of the two visual systems model is the claim that the human visual system can process one and the same stimulus in two fundamentally different ways. It can give rise to a visual percept (or to perceptual awareness), and to a visuomotor representation, of one and the same object, according to the task. The challenge is to provide an explicit characterization of the differences between visual perception and visually guided actions or between perceptual and visuomotor representations of a given visual stimulus.

Now, in healthy human adults, the two systems of visual processing beat to a single drum. So only via carefully designed experimental tasks can one system be dissociated from the other. The two visual systems model of human vision is first and foremost a view about the anatomy of the human visual system. In the second section of the paper, I will briefly recount the history of the two visual systems model of human vision. I believe that the two visual systems model of human vision runs counter to well-entrenched assumptions in the philosophy of perception, according to which the main purpose of human vision is visual perception. Thus, in the first section of the paper, I will lay out my five main assumptions, some of which will be supported by empirical evidence that I will selectively review in subsequent sections, some of which is inconsistent with what I take to be orthodox views in the philosophy of perception. In sections 3, 4 and 5, I will review experimental evidence respectively from electrophysiological studies in the monkey, neuropsychological studies of human patients and psychophysical studies in healthy subjects. Finally, in section 6, I will

argue that the perceptual awareness of the visual attributes of objects asymmetrically depends upon the representation of the spatial relations among two or more such objects.

1. Five background assumptions³

I am presently sitting on a chair in front of a table, facing an open window. On the table, I can see a tea-pot and a tea-cup next to it. I can see the shapes, contours, sizes, orientations, colors, textures of both the tea-pot and the tea-cup, and the level, the color and texture of the tea in the cup. The perception of the visual array containing the pair of objects on the table has a distinctive phenomenology: my visual percept makes me visually aware of a pair of objects with specific visual properties. The peculiar combination of visual properties bound in my percept is distinctive of the phenomenology of visual perception (by contrast with the phenomenology of experiences in other sensory modalities). Now, I can extend my right arm, reach and grasp the handle of the tea-cup with precision grip between my right thumb and index finger, bring the cup to my lips, and drink a sip of tea from it. It may seem as if my visually guided action of reaching for, and grasping, the handle of the tea-cup is under the control of my fully detailed visual percept representing the shapes, contours, sizes, orientations, colors and textures of both the tea-pot and the tea-cup in front of me. It may seem so, but it is not, or so I will argue.

1.1. First assumption

My first assumption (shared by many but not all philosophers) is that the detailed content of my visual percept outstrips my conceptual repertoire. I lack the detailed concepts and words in either French or English, whose content could match the content of my visual experiences of the shapes, contours, sizes, orientations, colors and textures of the pair of

³ In the course of stating my “assumptions”, I shall offer some justification for them.

objects in front of me. I will not rehearse but merely accept the arguments for the view that the conceptual content of my linguistically expressible thoughts and judgments fails to capture the fine-grainedness and the informational richness of the non-conceptual content of my visual percept.⁴

Conversely, there are things that one can think about but that one cannot see. For example, my visual percept can represent the cup as being to the right of the pot. It can represent the cup, the pot and their spatial relation. Now, I can form the thought that the cup is to the right of the pot from my present perspective, i.e., facing the open window. I can also form the thought that the cup is to the left (not to the right) of the pot for an observer sitting across the table from me with her back turned against the window. I can think that were I to sit across the table, I would see the cup to the left of the pot. But strictly speaking, I cannot see the cup as being to the right of the pot from my current perspective, let alone see the cup as being to the left of the pot from the opposite perspective. I cannot see the perspective from which (or by means of which) I can see the cup as being to the right of the pot while I occupy that very position. Of course, someone else can see the perspective in question. But I cannot. Granted, at $t + 1$, I can see both the cup to the left of the pot and the position I occupied at t after I have switched position from t to $t + 1$. But at t , when I see the cup to the right of the pot, I can think but I cannot see the position by means of which I see the cup to the right of the pot.

1.2. Second assumption

⁴ See e.g., Bermudez (1998), Dretske (1981), Evans (1982), Peacocke (1992, 1998, 2001), Tye (1995). For criticism, see McDowell (1994). Here, I merely state my view that demonstrative concepts fail to jointly satisfy requirements on concepts and to match the fine-grainedness and informational richness of visual percepts. If they match the fine-grainedness and informational richness of visual percepts, then they fail to satisfy conceptual requirements. If they satisfy conceptual requirements, then they fail to match the fine-grainedness and informational richness of visual percepts. See Jacob and Jeannerod (in press, ch. 1) for details.

Since (by the first assumption) the conceptual content of thoughts and linguistic utterances cannot match the fine-grainedness and informational richness of the non-conceptual content of visual percepts, many philosophers (e.g., O'Shaughnessy, 1992 and Peacocke, 1992) have linked the non-conceptual content of visual percepts to the requirements of bodily actions. In particular, they have assumed that the distinctive fine-grainedness and informational richness of the non-conceptual content of visual percepts are attuned to the on-line guidance of an agent's bodily movements.

According to the assumption which Clark (2001: 496) labels the assumption of 'experience based-control' (or EBC), the function of the detailed content of visual percepts is to guide and monitor bodily movements:

conscious visual experience presents the world in a richly textured way [...] that presents fine detail [...] (that may 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.

I reject EBC because it is, I think, inconsistent with the version of the two visual systems model of human vision to which I subscribe. I also reject O'Regan and Noë's (2001) version of the enactive theory of perceptual content, according to which the detailed pictorial content of a creature's visual experience is constituted by her implicit (or practical) knowledge (or mastery) of sensorimotor contingencies. As I shall argue (in section 4.3), the double dissociation between perceptual and visuomotor responses in brain-lesioned human patients, which is evidence for the two visual systems hypothesis, is also, on my view, evidence against the enactive theory.

Instead, I accept what Clark (2001) calls the assumption of 'experience-based selection' (EBS), according to which the fine-grainedness and informational richness of the non-conceptual content of visual percepts contribute to an agent's selection of the target of his

motor activity. In accordance with EBS, my second assumption is, in Goodale and Humphrey's (1998: 185) terms, that the non-conceptual content of visual percepts is not 'linked directly to specific motor outputs but [is] linked instead to cognitive systems involving memory, semantics, spatial reasoning, planning, and communication'. Thus, I assume that the non-conceptual content of visual percepts is poised for further conceptualization. In other words, visual percepts of objects provide the 'belief box' with visual information about objects. But a visual percept is not ipso facto a belief (or a judgment) with conceptual content.

1.3. Third assumption

The computational requirements made by the perceptual system on human vision are clearly different from the requirements made by the visual monitoring of object-oriented actions. The task of the former is to encode pictorial visual information about the enduring properties of objects that will facilitate identification or recognition from many different perspectives on many different occasions in which such things as the lighting, the distance from the objects, the relative motions of objects and the observer's own movements may vary considerably.⁵ Recognition of a perceived object requires that visual information matches conceptual information stored in long term memory. The task of the latter is both to encode the geometrical properties of the target relevant for prehension (such as size, shape and orientation) and to keep track of the spatial position of the target relative to the moving agent. As I will argue, it is of the essence of a visual percept that it represents objects in a visual array together with their spatial relations. Perceptual processing makes me visually aware of

⁵ By contrast with the conceptual content of thoughts, the content of visual percepts is pictorial in the sense that several visual attributes are always bound together in a visual percept. Philosophers sometimes misleadingly assume that enduring properties cannot be part of the content of representations with pictorial content. This is a mistake. Consider shape, which is uncontroversially an enduring property of an object. What makes a visual perceptual representation of shape pictorial is that it cannot represent the shape of an object without representing other properties (e.g., size and color) of the object. You can think of a triangle without thinking of its color or size. But in good lighting conditions, a healthy observer cannot visually perceive a triangular object unless he perceives its color, size and orientation. Nor can she visually image a triangular object without imaging its color, size and orientation.

the tea-cup, the tea-pot and their spatial relation, e.g., the former's being to the right of the latter. By contrast, for me to grasp the handle of the tea-cup and lift the cup off the table, my visual system must have computed two landing sites on the handle where to apply my right thumb and index finger. But I am not visually aware of these points. Furthermore, in order to reach the cup and grasp its handle, I must represent its location relative to my body, not relative to the tea-pot, unless the latter is on my way to the former.

According to my second assumption, conscious visual experience is involved in the selection of the target of an object-oriented action, not in the monitoring of the bodily action itself. According to my third assumption, what is crucial to the visual awareness of objects provided by visual perception is the representation of the spatial relations among constituent objects in a visual array. As we shall see, there is much *prima facie* paradoxical evidence that accurate visually guided actions does not require visual awareness of the target.

1.4. Fourth assumption

My fourth assumption is that much of human object-oriented behavior (such as pointing to, or reaching for and grasping, a target) is intentional. On the one hand, much of the visual behavior of humans and non-humans animals (such as the pupillary light reflex, the synchronization of circadian rhythms and the visual control of posture) is reflexive behavior, not intentional behavior. On the other hand, humans also perform many acts that are neither reflexive nor intentional, and which O'Shaughnessy (1980: 58-73) labels 'sub-intentional acts', such as the movement of one's tongue in one's mouth as one reads, the movements of one's fingers as one is attending to a philosophy lecture or the tapping of one's feet to the rhythm of a piece of music. Such sub-intentional acts are not reflexive behaviors: unlike the dilation of one's pupil, an agent can interrupt tapping her feet to the rhythm of the music if requested to do so. They are not intentional since the movements are not intended by the

agent, let alone produced by his or her beliefs and desires. By contrast, human movements of prehension are caused by the agent's intentions and they are monitored by a visuomotor representation of the target.

1.5. Fifth assumption

In accord with Anscombe (1957) and Searle (1983), I assume that perceptions and beliefs have a mind-to-world direction of fit. If their content matches a fact, then they are veridical. If it does not, then they are not. I assume that intentions and desires have a world-to-mind direction of fit. Their function is not to match a fact but to represent a goal, i.e., a non-actual state of affairs. In addition, percepts have a world-to-mind direction of causation: a percept is caused by the fact that it represents. Intentions have a mind-to-world direction of causation: an intention causes the obtaining of the state of affairs that it represents. Millikan (1996) has interestingly argued for the existence of Janus-like representations with both a mind-to-world and a world-to-mind direction of fit, which she calls 'pushmi-pullyu' representations. She argues convincingly that such representations with a twofold direction of fit are evolutionarily more primitive than representations with either a purely descriptive or a purely prescriptive direction of fit. My fifth assumption is that visuomotor representations are such hybrid mental representations with a twofold direction of fit. They represent what Gibson (1979) called affordances for manual actions. Thus, they provide motor intentions with visual information about the properties of objects relevant for actions of prehension.

2. A brief history of the two visual systems hypothesis

Early versions of the two visual systems hypothesis were first entertained by neurophysiologists working on the visual systems of non-human animals. In non-mammalian vertebrates, e.g., in amphibians, it was demonstrated by Ingle (1973) that prey-catching

behavior is mediated by retinal projections onto the optic tectum, while the visual control of barrier-avoidance is mediated by retinal projections onto pretectal nuclei. Similarly for mammals, it was demonstrated by Schneider (1967) that a hamster with a lesioned superior colliculus could discriminate vertical from horizontal stripes but could not run a maze. Conversely, a hamster with a lesioned visual cortex could run a maze but not do pattern recognition. Since the earlier evidence came from the study of animals with little or no visual cortex, early versions of the two visual systems hypothesis emphasized the contrast between on the one hand 'ambient' vision controlled by peripheral retinal information, based on subcortical structures, and on the other hand 'focal' vision based on cortical structures.

A first major step was taken by Ungerleider and Mishkin (1982), who located the two visual systems within the primate visual cortex. More than thirty specialized areas with many reciprocal connections have been discovered in the visual cortex of primates. But the connections turn out to be segregated into two main streams of processing: the ventral stream and the dorsal stream. The former projects the primary visual cortex (area V1) onto the inferotemporal area. The latter projects the primary visual cortex onto the posterior parietal lobe, itself a relay on the way towards the pre-motor cortex and the motor cortex (see Figure 1).

Ungerleider and Mishkin (1982) examined the selective effects of lesions in the brains of macaque monkeys on two kinds of behavioral tasks: a landmark task and an object-discrimination task. In the former task, the monkey had to discriminate between two covered wells — one empty and one full — according to whether they were located far away or near a landmark. In the latter task, the monkey had to discriminate two objects of different shapes, colors and textures. Ungerleider and Mishkin (1982) found that a lesion in the inferotemporal cortex severely impaired the animal in the object-discrimination task, but not in the landmark task. Conversely, they found that a lesion in the posterior parietal cortex severely affected the

animal's performance in the landmark task, but not in the object-discrimination task. On the basis of these experiments, Ungerleider and Mishkin (1982) concluded that the ventral stream is the What-system (or 'object-channel') and the dorsal stream is the Where-system (or 'space-channel'). Importantly, the landmark task tested the animal's ability to perceive spatial relations, not to act on a target. Thus, on Ungerleider and Mishkin's version of the two visual systems in the primate brain, the predominant function of both the ventral and the dorsal pathways is to mediate visual perception: while the former is involved in the perception of so-called 'intrinsic' attributes of an object such as its shape, size, texture and color, the latter is involved in the perception of such 'extrinsic' properties of objects as their spatial relations. On this view, visuomotor functions involved in guiding reaching and/or grasping movements are relegated to subcortical structures.

A second major step was taken by Goodale and Milner (1992) and Milner and Goodale (1995) when they provided room for the visuomotor transformation within their amended version of the two visual systems model of human vision.⁶ On their revised view, the ventral stream underlies what they call 'vision-for-perception' and the dorsal stream underlies what they call 'vision-for-action'. The crucial evidence on which Goodale and Milner based their revised interpretation of the two visual systems model of human vision is the neuropsychological double dissociation between two visual impairments produced by two selective lesions in the human visual system: whereas apperceptive visual agnosia results from lesions in the inferotemporal area, optic ataxia results from lesions in the posterior parietal cortex. Apperceptive visual agnosic patients are deeply impaired in the visual recognition of the size, shape and orientation of objects. But they can reach and grasp objects whose shapes, sizes and orientations they cannot visually recognize. Conversely, optic ataxic

⁶ The visuomotor transformation is the network in the human brain whereby visual information is converted into motor commands of arm and hand movement towards objects.

patients cannot reach and grasp objects whose shapes, sizes and orientations they can visually recognize.⁷

Thus, in the mid-1990's, the two major versions of the two visual systems model of human vision disagreed on the functional significance of the dorsal pathway and the role of the posterior parietal lobe. Ungerleider and Mishkin's (1982) model subscribes to the assumption that the major function of the primate visual system is to allow visual perception: all cortico-cortical connections in the primate visual brain underly perceptual awareness. By contrast, according to Milner and Goodale's (1995) model, neither is perceptual awareness the exclusive (or the main) function of vision in primates, nor are cortico-cortical pathways in the primate and the human brains limited to visual perception. Unlike Ungerleider and Mishkin (1982), Milner and Goodale (1995) locate the visuomotor transformation firmly within the visual cortex of non-human primates and humans.

3. Evidence for dual visual processing in the monkey

I will briefly review a sample of electrophysiological recordings of single cells in the brain of either anesthetized or awake monkeys that fall under two basic categories: cells in the ventral pathway of the monkey visual system are involved in the discrimination of basic visual attributes of complex stimuli such as their colors, textures, shapes and contours and cells in the posterior parietal cortex that respond to the geometric properties of objects that are relevant for tasks of prehension.

3.1. Perceptual cells in the inferotemporal area

Tanaka et al. (1991) recorded cells in the posterior part of the inferotemporal cortex of macaque monkeys that respond preferentially to relatively simple visual stimuli such as

⁷ See section 4 for detailed examination of the double dissociation between optic ataxia and apperceptive visual form agnosia.

elongated bars or small colored disks. Other cells in this area fire in response to the detection of some favored shape (e.g., an eight-arm star as opposed to a circle or a diamond). Still others fire in response to a complex combination of shape, orientation and texture (e.g., a triangle whose apex is directed towards the right with vertical stripes). In a more anterior part of the inferotemporal cortex, Tanaka (1993) found more elaborate cells that are sensitive to the contours of complex stimuli such as brushes, leaves and other biological objects. One such cell responded preferentially to the head of a tiger seen from the top. Tanaka (1993) found that the cell keeps responding to a series of simplifications of the stimulus until it is reduced to a combination of a pair of symmetrical black rectangles flanking a white square. It stops responding to either the pair of black rectangles or the white square alone (see Figure 2).

3.2. Visuomotor cells in the posterior parietal lobe

In the 1970's, Mountcastle et al. (1975) recorded cells in the posterior parietal cortex of awake macaque monkeys. They showed that, unlike inferotemporal cells, cells in the posterior parietal cortex covary with many aspects of the animal's visual behavior, such as saccadic eye-movements, head-movements and the monitoring of reaching towards, and manipulating, objects. Sakata et al. (1995) and Murata et al. (2000) recorded neurons in the anterior intraparietal (AIP) area of the awake monkey. They showed three dimensional black and white objects with well-demarcated geometrical properties: a plate, a ring, a cube, a cylinder, a cone and a sphere (see Figure 3). Four different conditions were compared, two of which involved an object-manipulation task (with or without a delay and with or without light), and two of which consisted of an object-fixation task (either in the presence or in the absence of light). AIP neurons turned out to divide into three groups: the visual dominant neurons fired preferentially when the animal manipulated the object in the light, not in the dark. The motor dominant neurons fired when the animal manipulated the object either in the

light or in the dark. The visual-and-motor dominant neurons fired more when the animal manipulated the object in the light than in the dark. Thus, AIP neurons are sensitive primarily to the geometric affordances of objects relevant for grasping and manipulating them.

Several experiments in the monkey have further showed that reaching and grasping can be selectively impaired. For example, Gallese et al. (1994) managed to cause a reversible impairment of the control of grasping that did not affect reaching. This experiment confirmed Mountcastle et al.'s (1975) distinction between 'arm projection' neurons and 'hand manipulation' neurons.

4. Double dissociations in brain-lesioned human patients

As indicated in section 2, much of Milner and Goodale's (1995) version of the two visual systems model of human vision relies on the double dissociation between apperceptive visual form agnosic patients and optic ataxic patients. Before examining this dissociation in more details, it is, however, necessary to review the psychophysical properties of movements of prehension in healthy human subjects.

4.1. Reaching and grasping in healthy human subjects

The articulation of their upper limbs and the dexterity of their hands allow non-human primates and especially humans to execute actions of prehension and the manipulation of objects. The human hand allows two basic kinds of grip: precision grip and power grip (see Jeannerod, 1997). The former is involved in grasping the handle of a tea-cup. The latter is involved in grasping a hammer. In the early 1980's, it was found that an action of prehension involves three major components or phases: a selection phase, a reaching phase and a grasping phase. During the selection phase, which is the perceptual part of the process, the target of prehension must be sorted out from a number of possible competitors and distractors

on perceptual grounds. The rest of the process is an automatic visuomotor process in which the visual information is exploited for the purpose of action.

The reaching part of the process of prehension involves ballistic movements of the different segments of the upper limbs connecting the hand to the shoulder. It was found that during transportation of the hand to the target, there is an automatic process of grip formation: the preshaping of the fingers is programmed much before the hand contacts the object. During transportation of the hand, first the finger-grip opens up and then closes down on the target. The opening reaches its peak — maximum grip aperture (or MGA) — at about 60% of the movement of transportation. It was found that the size of the finger-grip is both much larger than the size of the target and that it is linearly correlated with it. Calibration of the finger-grip does not require perceptual comparison of the hand with the target: it is made automatically on the basis of a visuomotor representation of the target without visual access to the hand (in so-called ‘open loop’ condition).

4. 2. Visuomotor impairments following lesions in the dorsal stream

Optic ataxia (also called Balint syndrome) is a disturbance in visually guided actions of pointing towards, and prehension of, objects in patients who have otherwise no purely motor impairment. It is subsequent upon a lesion generally located in the superior parietal lobe that can be located in either hemisphere and that causes impairment in movements of the hand contralateral to the lesion. The disturbance can either affect transportation of the hand towards the target or formation of the finger-grip appropriate for grasping or both (see Figure 4).

Patient AT (examined by Jeannerod et al. 1994, see also Jeannerod, 1997) presented a bilateral optic ataxia. She was impaired in the accuracy of both her reaching and grasping movements. In fact, she was far more severely impaired in grasping than reaching

movements. In particular, close examination revealed grossly exaggerated grip aperture. By contrast, her perceptual judgments were normal, as she could recognize and identify objects visually presented to her. Jeannerod et al. (1994) used the same effector in both a visuomotor task and a perceptual estimation task: the distance between her thumb and index finger. They found that while her grip aperture was severely inadequate in a task of grasping an object, the distance between her thumb and index finger was positively correlated with the size of objects in a task of perceptual estimation. Also when asked by Milner et al. (1999) to locate a source of light by pointing her index finger, her performance improved significantly when tested after a 5 second delay between the occurrence of the signal and the initiation of the action.

4.3. Perceptual impairments following lesions in the ventral stream

According to Farah (1990), there are two grades of visual form agnosia both caused by lesions in the ventral pathway of the human visual system: associative agnosia and apperceptive agnosia. The former, caused at a later stage of visual processing within the ventral stream, results, in Teuber's (1960) phrase, in a 'normal percept stripped of its meaning'.⁸ The patient can process the local shapes of objects but cannot identify and recognize them. The latter, caused at an earlier stage of visual processing within the ventral stream deprives the patient of the ability to perceive and thus be visually aware of the shapes, sizes and orientations of objects altogether.

Possibly, the distinction between the two grades of perceptual impairment in visual form agnosia could be matched onto Dretske's (1969, 1978) distinction between two distinct stages of normal visual perception: non-epistemic visual perception and primary epistemic visual perception. Suppose, for example, that you are driving too fast to identify some

⁸ Hence Jeannerod's (1994, 1997) contrast between the 'semantic' and the 'pragmatic' processing of visual information. In visual form agnosic patients, the semantic processing is impaired. In Jacob and Jeannerod (in press), we further distinguish a lower-level and a higher-level of pragmatic processing of objects. The former is involved in the visuomotor transformation and is impaired in optic ataxic patients. The latter is involved in the skilled manipulation of tools and is impaired in apraxic patients (with a lesion in the left inferior parietal lobe).

obstacle lying on your trajectory, but you skillfully manage to avoid hitting it. Unless you saw it, you would have hit it. Since you avoided it, you did see it. But since you failed to identify it, your perception of what you avoided was non-epistemic.⁹ Visual identification of an object yields perception of something as instantiating some property or other (or as falling under some concept or other). What Dretske (1969) calls ‘primary epistemic perception’ consists in seeing a fact involving a perceived object. For example, seeing that the cat is on the mat by seeing the cat is primary epistemic perception: unless you possess the concepts of a cat and of a mat, you could not see that the cat is on the mat.¹⁰ You could not form the belief that the cat is on the mat by visual means. Arguably, an apperceptive visual form agnostic could not come to believe that the cat is on the mat by seeing a cat lying on a mat: visually presented with a cat, she could not assemble its local shapes into the contour of a cat. In apperceptive visual form agnostic patients, earlier stages of non-epistemic perception of objects involving the assemblage of the elementary visual attributes into a percept of an object would be impaired. In associative visual form agnostic patients, later stages of primary epistemic perception of facts involving the matching between visual percepts and concepts of objects would be impaired.

It was an important finding to discover the residual visuomotor capacities of an apperceptive patient DF, first examined by Milner et al. (1991). This finding is reminiscent of the earlier finding of residual visuomotor capacities in blindsight patients after a lesion in their primary visual cortex (see Weiskrantz 1986, 1997). DF suffered a bilateral occipital lesion destroying a large part of her ventral pathway, as a consequence of carbon monoxide poisoning. When asked to name the shapes of geometrical objects and when asked whether two shapes were the same or different, DF was at chance. When asked to match the size of an

⁹ Here I am assuming that a driver is forming a visual percept, not a visuomotor representation, of an obstacle lying on the road on which she is driving.

¹⁰ In addition, what Dretske (1969, ch. III) calls ‘secondary’ epistemic visual perception is visually based belief about a fact involving an unperceived object. For example, seeing that the tank in your car is full by seeing not the tank but the gas gauge.

object by scaling the distance between the index finger and the thumb of her right hand, she was also at chance. When asked to report the orientation of a line (or a slot) or to match it by turning a hand-held card so as to match the perceived orientation of the line (or slot), her performance was very poor. DF's visual imagery was preserved: although she could not copy a visually presented object, she could draw one from memory. Finally, when asked to locate a source of light without a delay, DF's performance was normal and it deteriorated after a 2 second delay (see Goodale et al., 1991, Goodale et al., 1994, Goodale, 1995 and Milner and Goodale, 1995: 136-7).

By contrast with her deep perceptual impairment, DF was surprisingly accurate when the shape, size and orientation of an object had to be processed in the context of a goal-directed hand action. During reaching-and-grasping objects between her index finger and her thumb, DF was able to perform prehension movements with the very same objects whose shapes, sizes and orientations she could not visually recognize. While transporting a hand-held card in order to insert it into a slot at different orientations around the clock, she could normally turn her wrist and orient her hand (see Goodale et al., 1991, Goodale et al., 1994, Carey et al., 1996, Goodale, 1995, Milner and Goodale, 1995) (see Figure 5). DF's visuomotor capacities, however, turn out to have interesting perceptual limitations. For example, when asked to insert a T-shaped object into a T-shaped aperture, she had trouble matching the orientation of the stem with the orientation of the top. As noted by Goodale (1995: 197), it suggests that a dedicated system for fast and accurate visuomotor behavior (such as grabbing a branch) may need 'to coopt more flexible perceptual systems' when required to process the several axes of symmetry of a complex object.

The double dissociation between apperceptive visual form agnosia and optic ataxia is hard to square with O'Regan and Noë's (2001) and Noë's (2002) version of the enactive theory of perceptual content. On the one hand, optic ataxia shows that practical knowledge of

some basic sensorimotor contingencies is not a necessary condition of perceptual identification and recognition of visually presented objects. Although we may not be too sure about the phenomenology of her visual experience,¹¹ patient AT is quite able to correctly identify and recognize objects presented to her in the visual modality, but she is deeply impaired in visuomotor behaviors of reaching for, and grasping, objects. Impairment in reaching and grasping should presumably, according to the enactive theory, alter patient AT's ability to perceptually recognize visual objects. But it does not. On the other hand, apperceptive visual form agnosia shows that knowledge of some basic sensorimotor contingencies is not a sufficient condition for perceptual experience and recognition of visually presented objects. Patient DF has kept pretty remarkable visuomotor capacities. But she is nonetheless deprived of the perceptual experience of, and the ability to recognize the shapes, sizes and orientations of visually presented objects. If reaching and grasping an object do not constitute the sensorimotor contingencies knowledge of which should count for the perceptual experience of the object, I wonder which are the sensorimotor contingencies knowledge of which could underlie perceptual experiences.¹²

5. Dissociations between perception and action in healthy human subjects

Milner and Goodale's (1995) version of the two visual systems model of human vision was elaborated mostly on the basis of the double dissociation between optic ataxia and apperceptive visual form agnosia. The question arises: is the model consistent with what is known of the psychophysical responses of normal human subjects. This is the question addressed in the present section. I will start with dissociations between perceptual and

¹¹ As Kathleen Akins reminds me. In addition to her optic ataxia, patient AT also happens to have dorsal simultagnosia that impairs her recognition of more than one object at a time. But according to Milner et al. (1999), her simultagnosia is separable from her optic ataxia.

¹² I am reminded of the molecular biologist Jacques Monod's reaction to a claim made by the Piagetian psycholinguist Bärbel Inhelder that 'sensorimotor intelligence' is crucial to language acquisition, at the 1975 Royaumont Conference with Chomsky and Piaget. Monod observed that this claim could be tested by examining the acquisition of language in children with serious motor disabilities (see Piattelli-Palmarini, 1980). In response, Inhelder so weakened her claim that it became untestable.

visuomotor responses when the latter task consists in pointing towards a visible target with one's index finger. Pointing to a target involves the visual computation of the location of the target in 'egocentric coordinates', i.e., relative to the agent's body. Then, I shall examine dissociations between perceptual and visuomotor responses when the relevant task consists in grasping a target. Grasping an object involves processing the position of the object in egocentric coordinates and such attributes of the object as its shape, size and orientation.

5.1. Pointing to a target and perceiving it

Bridgeman et al. (1975) and Goodale et al. (1986) have exploited the phenomenon of saccadic suppression. When the motion of a target coincides with a saccadic eye-movement, a healthy human subject is not consciously aware of the motion of the target. Bridgeman et al. (1979) and Goodale et al. (1986) found that normal human subjects can point accurately to a target on the screen of a computer whose motion they could not consciously notice because it coincided with one of their saccadic eye-movements (see Bridgeman, 2002). Castiello et al. (1991) found that subjects are able to correct the trajectory of their hand movement directed towards a moving target some 300 milliseconds before they became conscious of the target's change of location (see Jeannerod, 1997: 82).

Pisella et al. (2000) and Rossetti and Pisella (2000) performed experiments involving a pointing movement towards a target. Subjects were presented with a green target towards which some of them were requested to stop their pointing movement when and only when they saw the target change location by jumping either to the left or to the right. Among the responses of the subjects whose task was to stop pointing, they found a significant percentage of very fast unwilled correction movements towards the target in response to a change of location to the left or to the right. Pisella et al. (2000) and Rossetti & Pisella (2000) called the mechanism responsible for these fast unwilled correction movements the 'automatic pilot' for

hand movement. In a second experiment, Pisella et al. (2000) presented subjects simultaneously with pairs of a green and a red target. They were instructed to point to the green target, but the color of the two targets could be interchanged unexpectedly at movement onset. Unlike a change of target location, a change of color did not elicit fast unwilled corrective movements by the automatic pilot. Unlike motion, which is processed by brain area MT, which belongs to the dorsal pathway, color is processed in areas that are more ventrally located. On this basis, Pisella et al. (2000) draw a contrast between the fast visuomotor processing of the location of a target in egocentric coordinates and the slower visual processing of the color of an object. In an optic ataxic patient with a lesion in the superior parietal lobe, the fast automatic corrections for changes of target-location were missing. The authors conclude that the automatic pilot that produces fast corrections of hand movements by normal subjects in response to target-motion is located in the superior parietal cortex.

In an experiment by Gentilucci et al. (1996), subjects sat in front of a table on which they could see three possible stimuli: one of the two segments of the Müller-Lyer size-contrast illusion with opposing arrow configurations or a control segment without arrows (see Figure 6). Subjects positioned their right index finger towards the nearest vertex of the displayed segment. They were asked to point their index finger from its initial position towards the more distant vertex of the segment. Gentilucci et al. (1996) compared four conditions:

- in the ‘full-vision’ condition, subjects could see both their hand and the target during the action;
- in the ‘non-visual feedback condition’, they could see the target, but not their hand;
- in the ‘no-vision’ condition, they could see neither their hand nor the target during the action;

— in the ‘no-vision 5 second delay’ condition, they could see neither their hand nor the target during the action, which started 5 seconds after the light went off.

Gentilucci et al. (1996) found an increasing effect of the Müller-Lyer illusion on the pointing movement from the first to the fourth condition. When subjects can see both their hand and the target, they code its location within a visuomotor representation of the target, in which the target’s location is coded in egocentric coordinates, relative to either their body or their hand. When subjects build such a visuomotor representation of the target, they do not need to represent the full segment at all. The target is just a point in egocentric space where to direct one’s index finger. A fortiori, the arrow configurations surrounding the vertices of a Müller-Lyer segment are not part of the visuomotor representation of the target. By contrast, when subjects see neither their hand nor the target, they code the location of the target as the furthest vertex of a segment. Thus, they build a perceptual representation of the length of the segment and the orientations of the arrows become relevant.

5.2. Grasping an illusory display

In standard displays of the Titchener (or Ebbinghaus) size-contrast illusion, two circles of equal diameters are surrounded by two annuli of circles either smaller or larger than the central circle. The central circle surrounded by an annulus of circles smaller than it looks larger than the central circle surrounded by an annulus of circles larger than it (see Figure 7). Pairs of unequal circles can also be made to look equal. Aglioti et al. (1995) replaced the 2-D central circles by 3-D graspable disks. In a first row of experiments with pairs of unequal disks whose diameters ranged from 27 mm to 33 mm, they found that on average the disk in the annulus of larger circles had to be 2,5 mm wider than the disk in the annulus of smaller circles in order for both to look equal. Aglioti et al. (1995) alternated presentations of physically unequal disks, which looked equal, and presentations of physically equal disks,

which looked unequal. Both kinds of trials were presented randomly and so were the left vs. right positions of either kind of stimuli. Subjects were instructed to pick up the disk on the left between the thumb and index finger of their right hand if they thought the two disks to be equal or to pick up the disk on the right if they judged them to be unequal (see Figure 8).

The sequence of subjects' choices of the disk on the right vs. left provided a measure of the magnitude of the illusion prompted by the perceptual comparison between two disks surrounded by two distinct annuli. In the visuomotor task, the measure of grip size was based on the unfolding of the natural grasping movement performed by subjects while their hand approached the object. During a prehension movement, fingers progressively stretch to a maximal grip aperture (MGA) before they close down until contact with the object (see section 4.1.). Aglioti et al. (1995) measured MGA in flight using optoelectronic recording. They found that grip aperture was significantly less affected by the size-contrast illusion than comparative perceptual judgment as expressed by the sequence of choices of disks on the left vs. right.

This experiment, however, raises a number of methodological problems. The main issue, raised by Pavani et al. (1999) and Franz et al. (2000), is the asymmetry between the two tasks. In the perceptual task, subjects are asked to compare two distinct disks surrounded by two different annuli. But in the grasping task, subjects focus on a single disk surrounded by an annulus. So the question arises whether, from the observation that the comparative perceptual judgment is more affected by the illusion than the grasping task, one may conclude that perception and action are based on two distinct representational systems.

Aware of this problem, Haffenden & Goodale (1998) performed the same experiment, but they designed one more task: in addition to instructing subjects to pick up the disk on the left if they judged the two disks to be equal in size or to pick up the disk on the right if they judged them to be unequal, they also required subjects to manually estimate between the

thumb and index finger of their right hand the size of the disk on the left if they judged the disks to be equal in size and to manually estimate the size of the disk on the right if they judged them to be unequal. Haffenden & Goodale (1998) found that the effect of the illusion on the manual estimation of the size of a disk (after comparison) was intermediary between comparative judgment and grasping.

Furthermore, Haffenden & Goodale (1998) found that the presence of an annulus had a selective effect on grasping. They contrasted the presentation of pairs of disks either against a blank background or surrounded by an annulus of circles of intermediate size, i.e., of size intermediary between the size of the smaller circles and the size of the larger circles involved in the contrasting pair of illusory annuli. The circles of intermediate size in the annulus were slightly larger than the disks of equal size. When a pair of physically different disks were presented against either a blank background or a pair of annuli made of intermediate size circles, both grip scaling and manual estimates reflected the physical difference in size between the disks. When physically equal disks were displayed against either a blank background or a pair of annuli made of circles of intermediate size, no significant difference was found between grasping and manual estimate. The following dissociation, however, turned up: when physically equal disks were presented with a middle-sized annulus, overall MGA was smaller than when physically equal disks were presented against a blank background. Thus, the presence of an annulus of middle-sized circles prompted a smaller MGA than a blank background. Conversely, overall manual estimate was larger when physically equal disks were presented against a background with a middle-sized annulus than when they were presented against a blank background. The illusory effect of the middle-size annulus presumably arises from the fact that the circles in the annulus were slightly larger than the equal disks. Thus, whereas the presence of a middle-sized annulus contributes to increasing manual estimation, it contributes to decreasing grip scaling. This dissociation

shows that the presence of an annulus may have conflicting effects on perceptual estimate and on grip aperture.

Finally, Haffenden et al. (2001) went one step further. They presented subjects with three distinct Titchener circle displays one at a time, two of which are the traditional Titchener central disk surrounded by an annulus of circles either smaller than it or larger than it. In the former case, the gap between the edge of the disk and the annulus is 3 mm. In the latter case, the gap between the edge of the disk and the annulus is 11 mm. In the third display, the annulus is made of small circles (of the same size as in the first display), but the gap between the edge of the disk and the annulus is 11 mm (like the gap in the second display with an annulus of larger circles) (see Figure 9). What Haffenden et al. (2001) found was the following dissociation: in the perceptual task, subjects estimated the third display very much like the first display and unlike the second display. In the visuomotor task, subjects' grasping in the third condition was much more similar to grasping in the second than in the first condition (see Figure 10). Thus, perceptual estimate was far more sensitive to the size of the circles in the annulus than to the distance between target and annulus. Conversely, grasping was far more sensitive to the distance between target and annulus than to the size of the circles in the annulus. The idea here is that the annulus is processed by the visuomotor processing as a potential obstacle for the position of the fingers on the target disk. If so, then the visuomotor processing of objects can be fooled: it can misrepresent (and thus represent) a 2D annulus as a 3D obstacle.

In Haffenden et al.'s (2001) experiment, two parameters are relevant: the size of the diameter of the circles in the annulus and the distance between the disk and the annulus. The suggested conclusion of their experiment is that what matters to the illusory visual percept is the contrast between the size of the disk and the size of the circles in the annulus. What matters to the visuomotor representation of the same display is the distance between the target

and the annulus. But Haffenden et al. (2001) consider only three conditions, not four. The question arises: does the two visual systems model lead to the prediction that the distance (between target and annulus) should produce the same dissociation between perceptual estimate and MGA when the disk is surrounded by larger circles? Given the difference in the size of the whole display, it is clear, I think, that the two visual systems model is not committed to this prediction. When the circles in the annulus and the distance between the disk and the annulus are smallest, then the size of the whole display is 56 mm. When the circles in the annulus and the distance between the disk and the annulus are largest, then the size of the whole display is 160 mm. If the circles in the annulus were large and the distance between the disk and the annulus were small, the size of the whole display would be 146 mm. What is true of a visuomotor response in the presence of a 56 mm display is not automatically true of the visuomotor response to a 146 mm display. Suppose that if the gap between the target and the potential obstacle is small, then the visuomotor programs automatically grasping of a 56 mm display. It does not entail that one produces the same visuomotor response to a 146 mm display.

Taken seriously, Haffenden et al.'s (2001) evidence shows that it is wrong to assume, as e.g., Milner and Goodale (1995) and Haffenden and Goodale (1998) have, that unlike the perceptual processing, the visuomotor processing of a visual display cannot 'afford the luxury' of being fooled by a visual display. Rather, the features of a display that can fool a visuomotor representations are not the same as those that can fool a perceptual representation. Automatic computation of the contrast between the sizes of elements of the display gives to rise to a perceptual size-contrast illusion. Automatic computation of the distance between a 3D target and 2D surrounding elements can give rise to a visuomotor illusion (representing the 2D elements as if they were 3D).¹³ If so, then presumably visuomotor processing

¹³ We normally grasp 3D targets against a 3D background. So it makes sense, as Kathleen Akins suggests, to assume that representing a 3D background is the default setting of the visuomotor representation of a target for

generates genuine visuomotor representations. Properly analyzed, the psychophysical data point, I think, away from what Franz et al. (2001) call the ‘common representation’ model of human vision, and towards the ‘separate representations’ model, according to which selective tasks can prompt the elaboration of distinct visual representations of one and the same visual stimulus.

6. Spatial processing and visual awareness

One can think about, but one cannot see, entities that are not in space (e.g., numbers). Arguably, the visual system cannot represent an object perceptually or otherwise unless it locates the object in space. As preceding sections have made clear, the computational requirements made respectively by the perceptual system and by the visuomotor system upon the visual encoding of spatial information about objects are clearly different. There are at least two broad ways the location of an object can be visually coded: it can be coded in egocentric and in allocentric coordinates. I will presently review evidence in support of the third of my five assumptions: representing the spatial relations among objects in a visual scene is necessary for perceptual awareness of the other visual attributes of objects. By contrast, in a visuomotor representation, the position of the target (of a hand action) is coded in an egocentric frame of reference and it does not lead to perceptual awareness of the object’s other visual attributes.

6.1. Perceptual proto-objects

In an experiment based on a habituation/dishabituation paradigm designed by Wynn (1992a, 1992b), 4.5 month old human infants see a human hand bring one Mickey Mouse

action. Still, representing a 2D annulus as a 3D obstacle is misrepresenting it. What is important for my purpose is that an agent, who automatically misrepresents a 2D annulus as a 3D obstacle as part of her visuomotor representation of a target, may of course become perceptually aware that the annulus is a 2D display. But then she must switch from a visuomotor illusory representation to a correct perceptual representation of the annulus.

onto a stage. Then, the puppet is hidden behind a removable screen. The babies see the empty hand leave the stage. Then, the babies see the hand bring a second puppet onto the stage with the screen up so that the babies do not see what is behind. Then, the babies see the empty hand leave the stage one more time. When the screen is removed, the babies can see either two puppets or one. By measuring the babies' looking time, Wynn found that they look longer at one than at two puppets. The experiment shows that 4.5 month old human babies can distinguish between two objects on the basis of spatial and locational information, i.e., from seeing twice a hand that holds an object and then leaves empty.

In a set of experiments using the same habituation/dishabituation paradigm, Carey (1995) and Xu and Carey (1996) showed 10 month old human infants alternatively a red metal car and a brown teddy bear. The babies saw the red metal car appear from, and disappear behind, the left side of a screen. They saw the brown teddy bear appear from, and disappear behind, the right side of the screen. They never saw the two objects next to each other. When the screen was removed, the infants saw either two or only one object. Measuring the infants' looking time, Carey (1995) and Xu and Carey (1996) found that 10 month old human babies were not more surprised to see only one object rather than two. They conclude that, in the absence of visual spatial locational information, 10 month old human infants fail to infer that they are being presented with two distinct objects on the basis of visual information about the colors, shapes, size and texture of objects. Featural information about colors, shapes, size and texture of pairs of objects presented separately was not sufficient to warrant in 10 month old human infants the judgment that they are presented with two distinct objects.¹⁴

Leslie et al. (1998) argue further that objects can be visually 'indexed' by two distinct indexing mechanisms: object-individuation and object-identification. They claim that the

¹⁴ Experiments by Bonatti et al. (2002) suggest that 10 month old human infants respond differently if presented with human faces or human-like stimuli.

dorsal pathway of the human visual system is the anatomical basis of the object-individuation mechanism that works by locational information. They claim that the ventral pathway is the anatomical basis of the object-identification mechanism that uses featural information (such as color, shape or texture). Finally, they argue that 10 month old human infants can only make use of the object-individuation mechanism, not of the object-identification mechanism. If the experiments are to be trusted, then 10 month old human infants use locational information, not featural information, in perceptual object-discrimination tasks.¹⁵

Pylyshyn (2000a, 2000b) reports relevant experiments about so-called ‘multiple object-tracking’ (MOT) in normal human adults. Subjects are shown eight identical circles at rest on a screen, four of which flicker briefly. Then, subjects see the eight circles move randomly on the screen for about 10 seconds. They are asked to keep track of the four circles that initially flickered. Normal human adults can keep track of four-to-five such distinct objects (or proto-objects). Importantly, Pylyshyn reports that subjects failed to notice changes in the colors and shapes of the proto-objects that they tracked by their relative locations.

Presumably, in MOT experiments, human adults use locational information, not featural information, to individuate and keep track of four-to-five randomly moving circles. At least, when a conflict arises between information about an object’s location and featural information, locational information trumps information about the object’s other visual properties. Now, in order to keep track of three to four proto-objects by locational information requires that the relative spatial positions of each proto-object be coded in an allocentric frame of reference centered on one of them. The very fact that adults engaged in MOT tasks fail to notice changes in colors, textures, shapes and other visual features of objects suggests that the object-individuation mechanism (in Leslie et al.’s 1998 sense) — or the system for representing spatial relations among proto-objects — must be a modular (informationally

¹⁵ There is, however, conflicting evidence that 4.5 month old infants’ perceptual segregation of objects from the background can benefit from prior exposure to the object (see e.g., Needham, 2001). So my claim about the role of locational information in infants’ perception of proto-objects is conditional.

encapsulated) system designed for the visual perception of spatial relations among proto-objects. This modular system, which represents the spatial positions of proto-objects in allocentric coordinates, is impervious to the perception of colors and other featural visual attributes of objects.

6.2. Unilateral spatial neglect

Lesions located in the right inferior parietal lobe typically produce unilateral spatial neglect. Unlike lesions in the superior parietal lobe, which produce optic ataxia and which can be bilateral, lesions responsible for unilateral spatial neglect are located in the right hemisphere. Patients with unilateral spatial neglect are not perceptually aware of objects visually presented in their contralesional (i.e., left) hemispace. For example, when asked to mark line segments at different orientations, a neglect patient will systematically fail to mark the segments lying in her contralesional hemifield. If asked to bisect a horizontal line, she will exhibit a strong ipsilesional bias revealing neglect of the part of the line falling within her neglected hemispace (see Milner et al., 1993).

Unlike blindsight patients whose primary visual cortex has been damaged, and to a lesser extent unlike visual form agnostic patients whose ventral stream has been impaired, neglect patients lack perceptual awareness on their affected side in spite of the fact that the visual pathway for processing the neglected visual information remains intact. Indeed, there is considerable evidence for covert processing of the neglected stimuli. For example, Marshall and Halligan (1994) showed neglect patient PS drawings of two houses located on top of each other, one of which displayed brightly colored flames on its left side. When asked to make an explicit comparison between the two houses, the patient could report no difference. When asked, however, which of the two houses she would rather live in, the patient pointed to the

house without flames. This shows that the neglected stimuli are covertly processed in neglect patients even though this processing is not accompanied by perceptual awareness.

One important source of insight into neglect are studies of a phenomenon called ‘extinction’, reported by Driver and Vuilleumier (2001). Unlike blindsight patients, who typically will fail to notice the presence of a light presented in isolation to their left, patients with unilateral neglect in their contralesional left hemispace might detect an isolated stimulus on their left easily. If, however, they are presented with two competing stimuli, one further to the left than the other, then they will typically fail to perceive the one further to the left. Thus, extinction is the failure to perceive a stimulus presented further within the contralesional hemispace, when the patient does perceive a competing stimulus presented more towards the ipsilesional hemispace. The stimulus located more towards the ipsilesional side of the lesion ‘extinguishes’ its competitor located more towards the contralesional side. As Driver and Vuilleumier (2001) emphasize, extinction reveals that neglect patients have a deep impairment in allocating attentional resources to competing stimuli according to their respective positions in the patient’s hemispace.

Driver and Vuilleumier (2001: 52-54) report the following extinction study with a neglect patient. In some conditions, the stimulus was a Kanizsa white square whose subjective contours resulted from the removal of the relevant quarter-segments from four black circles appropriately located. In other conditions, the stimulus consisted of the four black circles in the same spatial position from which the same relevant quarter-segments had been removed, except for the persistence of narrow black arcs at the outer boundary of the removed quarter-segments that prevented the formation of the subjective contours of the Kanizsa white square (see Figure 11). The patient extinguished most left-sided presentations of the stimulus in bilateral trials when the narrow black arcs prevented the formation of the subjective contours of the Kanizsa white square. But extinction decreased significantly upon the perception of the

Kanizsa square. Thus, neglect patients find it easier to allocate their perceptual attention to one object (the white Kanizsa square) than to four competing distinct objects (the four black circles each with a partially removed quarter-segment) with four distinct relative spatial positions.¹⁶

Driver and Vuilleumier (2001: 54) further report a remarkable modulation of extinction according to the task. When asked to report the location of a stimulus (e.g., a small set of distinctive shapes) or when asked whether they perceive a small set of distinctive shapes at a given location, neglect patients extinguish left-sided stimuli in bilateral displays. However, changing the task decreases the extinction. If the task is to enumerate the members of the very same small set of shapes in bilateral displays, neglect patients succeed. For small numerosities (i.e., sets of three or perhaps even four objects), humans can provide precise numerical estimation by subitizing. The ability to subitize small numerosities is the ability to produce direct perceptual estimates of small numerosities. It seems to rely on the ability to build a visual representation of the distinct relative spatial positions of objects within a set (see Dehaene, 1997, ch. 3). Arguably, one condition for exploiting a subitizing procedure is to form a perceptual representation of a single set of small cardinality. If so, then the contrast between the two tasks is the contrast between attending to e.g., three competing distinct entities as opposed to attending to a single object involving three distinct elements, parts or components. In the first task, the stimuli compete for the patients' perceptual attention in the neglected hemispace and competition produces extinction. In the second task, the very same stimuli are transformed into elements of a single object with no competitor.

6.3. Awareness respectively of space and of other visual properties

¹⁶ There is evidence that activity in area V2 is necessary for perceiving the subjective contours of a Kanizsa figure (see Zeki, 1993: 316-7). If so, then area V2 is active in neglect patients.

Joint work of Wynn (1992a, 199b), Carey (1995), Xu and Carey (1996) and Leslie et al. (1998) suggests very strongly that 4.5 month old human infants automatically open different object files on the basis of their perception of the relative locations of objects. If 10 month old human infants are not provided perceptual information about the distinct locations of two distinct objects, then featural information alone (e.g., color, shape, orientation and texture) is not sufficient to warrant the opening of two distinct object-files. I call ‘proto-objects’ such object-files opened on the basis of relative locations. I think infants’ proto-objects are quite similar to what human adults manage to track in Pylyshyn’s MOT tasks. Subjects of such experiments manage to attend to the random motions of four-to-five proto-objects identified by their relative positions on a screen, not by their colors, shapes or texture.

Lesions in the right inferior parietal lobe produce impairment in the ability to perceive spatial relations among objects in the patient’s left contralesional hemispace. Driver and Vuilleumier (2001) adduce much evidence that in neglect patients, the neglected stimuli are actively processed in the brain areas (particularly, in the ventral pathway) that, in healthy human subjects, give rise to conscious visual experiences of objects’ shapes, sizes, colors, texture, and so on. As Block (2001: 198) rightly emphasizes, the condition of neglect patients is paradoxical: the brain areas (in the relevant portions of the ventral pathway) that should make neglect patients aware of the visual properties of extinguished stimuli in their contralesional side are active, but the patient remains unaware of these stimuli.

Driver and Vuilleumier (2001), however, suggest a solution to the paradox: the solution is the ‘neuropsychological asymmetry’ between awareness of the relative locations and awareness of other properties of visual stimuli. By losing visual awareness of the relative spatial locations of visual stimuli in their neglected side, neglect patients also lose visual awareness of their colors, sizes, shapes, orientations and textures. Although the properties of neglected stimuli are unconsciously processed in the relevant brain areas of neglect patients,

nonetheless these patients remain visually unaware of them. In Driver and Vuilleumier's (2001: 74) terms, 'it thus appears that when the appropriate representation of a stimulus location is lost or degraded, as in neglect after parietal damage, then awareness of other stimulus properties (presumably coded elsewhere in the brain) is also lost'. As Driver and Vuilleumier (2001: 75) note, the converse is not true: loss of awareness of such visual properties of stimuli as their colors, shapes, sizes or orientations does not lead to unawareness of the relative locations of stimuli.¹⁷

Concluding remarks on the two visual systems hypothesis

Human optic ataxic patients have a possibly bilateral lesion in the superior parietal lobe. As a result, their visuomotor transformation is deeply altered: they cannot properly represent the size, shape and orientation of a target in a visuomotor representation in which its location is coded in egocentric coordinates centered on the patient's body. They can, however, form appropriate visual percepts of objects by means of which they can recognize the shapes, sizes and orientations of objects. Apperceptive visual form agnostic patients have a lesion in the inferotemporal areas (in the ventral pathway). As a result, they lose visual awareness of the shapes, sizes and orientations of objects. They can, however, accurately compute the size, shape and orientation of a target for the purpose of a visuomotor representation in which the location of the target is coded in egocentric coordinates. Neglect patients have a lesion in the right inferior parietal lobe but the rest of their visual cortex is intact. As a result of their inability to represent the respective spatial locations of objects in their neglected hemispace, they fail to reach awareness of the other visual attributes of these objects. As Driver and Vuilleumier (2001: 70) write:

¹⁷ Dennett's (2001) emphasis on the role of competitive mechanisms in producing perceptual awareness is, it seems to me, on the right track. But he fails to notice the asymmetry between awareness of spatial relationships among objects and awareness of other visual attributes of objects.

The pathology in [neglect] patients becomes more evident when several stimuli are presented simultaneously [...] Interestingly, extinction in the patients can be determined less by the absolute location of a stimulus within the left or right hemifield than by its position relative to other competing stimuli.

Thus, such patients should not have serious trouble in reaching for one stimulus in their neglected hemispace, since they might be able to form a visuomotor representation of a target of prehension whose 'absolute' location is coded in egocentric coordinates. Indeed, there is evidence that neglect patients have such visuomotor capacities as reaching and grasping objects located in their left hemispace.

The neuropsychological picture suggests that, from an anatomical point of view, the parietal lobe (which is part of the dorsal pathway of the human visual system) makes a twofold contribution to human vision: the superior parietal lobe makes a contribution to the visuomotor transformation. It allows the formation of a visuomotor representation of a target whose location is coded in egocentric coordinates centered on the agent's body. This system is still functioning in apperceptive visual form agnosic patient DF. The right inferior parietal lobe makes a decisive contribution to conscious visual perception: it allows the representation of the location of an object within allocentric coordinates centered on some item in the visual scene. Neglect patients show that unless the location of an object can be coded in allocentric coordinates, the object's other visual attributes are not available for visual awareness.

On the one hand, as apperceptive visual form agnosic patient DF shows, unless the relevant areas in the ventral pathway are active, processing of shape, size and orientation of an object will not yield visual awareness of shape, size and orientation of an object. In particular, if such attributes are only processed for building a visuomotor representation of an object whose location is coded in egocentric coordinates, then conscious awareness fails to arise. On the other hand, as neglect patients show, active processing in the relevant areas in the ventral

pathway for the shape, size and orientation of an object is not sufficient for visual awareness of the object's shape, size and orientation either. Representing the relative spatial locations of an object in allocentric coordinates is a necessary condition for awareness of the object's shape, size and orientation. Milner and Goodale (1995: 200) are right to claim that 'visual phenomenology [...] can only arise from processing in the ventral stream, processing that we have linked with recognition and perception'. But what they say may wrongly suggest that processing in the ventral stream is sufficient for visual phenomenology — which it is not. Block (2001: 199) asks the question: 'what is the missing ingredient, X, which, added to ventral activation (of sufficient strength), constitutes conscious [visual] experience?'. To Block's question, I venture to answer that, for many of an object's visual attributes, X is: the coding of the object's spatial location within an allocentric frame of reference by the right inferior parietal lobe.¹⁸

In Jacob and Jeannerod (in press), we claim that for awareness of such visual attributes of an object as its shape, size and orientation to arise, the visual representation must satisfy what we call the constraint of contrastive identification. The visual processing of an object's shape, size and orientation must be available for comparison with the shapes, sizes and orientations of other neighboring constituents of the visual array. It is of the essence of conscious visual perception that it makes automatic comparisons, which are the source of e.g., size-contrast perceptual illusions. Only if the spatial location of an object is coded in allocentric coordinates will the processing of the object's other visual attributes be available for visual awareness. This is exactly what is missing in a purely visuomotor representation of

¹⁸ For a slightly different view of the contribution of the parietal lobe to visual awareness, see Kanwischer (2001: 108) who writes: 'neural activity in specific regions within the ventral pathway is apparently correlated with the content of perceptual awareness, whereas neural activity in the dorsal pathway may be correlated instead with the occurrence of perceptual awareness in a completely content-independent fashion'. If awareness of spatial relationships among objects counts as 'completely content-independent perceptual awareness', then I think that I agree with Kanwischer (2001).

a target of prehension whose ('absolute') location is coded in egocentric coordinates. Table 1 summarizes the leading ideas:

	world-to-mind direction of fit	mind-to-world direction of fit	localization of target in egocentric coordinates	localization of object in allocentric coordinates	awareness of visual attributes
visuomotor representation	+	+	+	-	-
visual percept	-	+	-	+	+

Table 1

References

Anscombe, G.E. (1957) Intention, Oxford: Blackwell.

Austin, J. (1962), Sense and Sensibilia, Oxford: Oxford University Press.

Bermudez, J. (1997) The Paradox of Self-Consciousness, Cambridge, Mass.: MIT Press.

Block, N. (2001) Paradox and cross purposes in recent work on consciousness, in Dehaene, S. & Naccache, L. (eds.) The Cognitive Neuroscience of Consciousness, Cambridge, Mass.: MIT Press.

Boussaoud, D., Ungerleider, L. & Desimone, R. (1990) Pathways for motion analysis: cortical connections of the medial superior temporal sulcus and fundus of the superior temporal visual areas in the macaque monkey, Journal of Comparative Neurology, 296, 462-495.

Bridgeman, B., Lewis, S., Heit, G. & Nagle, M. (1979) Relation between cognitive and motor-oriented systems of visual position perception, Journal of Experimental Psychology: Human Perception and Performance, 5, 692-700.

Bridgeman, B. (2002) Attention and visually guided behavior in distinct systems, in Prinz, W. & Hommel, B. (eds) Common Mechanisms in Perception and Action, Oxford University Press.

Carey, D.P., Harvey, M. & Milner, A.D. (1996) Visuomotor sensitivity for shape and orientation in a patient with visual form agnosia. Neuropsychologia, 34, 329-337.

Carey, S. (1995) "Continuity and Discontinuity in Cognitive Development", in Osherson, D. (ed.) (1995) An Invitation to Cognitive Science, Thinking, vol. 3, Cambridge, Mass.: MIT Press.

Clark, A. (2001) Visual experience and motor action: are the bonds too tight?, Philosophical Review, 110, 495-519.

Dehaene, S. (1997) The Number Sense: How the Mind Creates Mathematics, New York: Oxford University Press.

Dennett, D. (2001) Are we explaining consciousness yet?, in Dehaene, S. & Naccache, L. (eds.) The Cognitive Neuroscience of Consciousness, Cambridge, Mass.: MIT Press.

Dretske, F. (1969) Seeing and Knowing, Chicago: Chicago University Press.

Dretske, F. (1978) "Simple seeing", in Dretske, F. (2000) Perception, Knowledge and Belief, Cambridge: Cambridge University Press.

Dretske, F. (1981) Knowledge and the Flow of Information, Cambridge, Mass.: MIT Press.

Driver, J. & Vuilleumier (2001) Perceptual awareness and its loss to unilateral neglect and extinction, in Dehaene, S. & Naccache, L. (eds.) The Cognitive Neuroscience of Consciousness, Cambridge, Mass.: MIT Press.

Evans, G. (1982) The Varieties of Reference, Oxford: Oxford University Press.

Farah, M.J. (1990) Visual agnosia. Disorders of Object Recognition and What They Tell Us About Normal Vision. Cambridge (Mass), MIT Press.

Franz, V.H., Fahle, M., Bühlhoff, H.H. & Gegenfurtner, K.R. (2001) Effects of visual illusion on grasping, Journal of Experimental Psychology: Human Perception and Performance, 27, 1124-1144.

Franz, V.H., Gegenfurtner, K.R., Bühlhoff, H.H. & Fahle, M. (2000) Grasping visual illusions: no evidence for a dissociation between perception and action, Psychological Science, 11, 20-25.

Gallese, V., Murata, A., Kaseda, M., Niki, N. & Sakata, H. (1994) Deficit of hand preshaping after muscimol injection in monkey parietal cortex, Neuroreport, 5, 1525-1529.

Gibson, J.J. (1979) The Ecological Approach to Visual Perception, Boston: Houghton-Mifflin.

Goodale, M. A. (1995) The cortical organization of visual perception and visuomotor control, in Osherson, D. (1995)(ed.) An Invitation to Cognitive Science, Visual Cognition, vol. 2, Cambridge, Mass.: MIT Press.

Goodale, M. and Humphrey, G.K. (1998) The objects of action and perception, Cognition, 67, 191-207.

Goodale, M.A., Jakobson, L.S., Milner, A.D., Perrett, D.I., Benson, P.J. & Hietanen, J.K. (1994) The nature and limits of orientation and pattern processing supporting visuomotor control in a visual form agnostic, Journal of Cognitive Neuroscience, 6, 46-56.

Goodale, M.A. & Milner, A.D. (1992) Separate visual pathways for perception and action, Trends in Neuroscience, 15, 20-5.

Goodale, M. A., Milner, A.D., Jakobson I.S. and Carey, D.P. (1991) A Neurological dissociation between perceiving objects and grasping them, Nature, 349, 154-56.

Goodale, M.A., Pélisson, D. & Prablanc, C. (1986) Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement, Nature, 320, 748-750.

Haffenden, A. M. & Goodale, M. (1998) The effect of pictorial illusion on prehension and perception, Journal of Cognitive Neuroscience, 10, 1, 122-36.

Haffenden, A.M. Schiff, K.C. & Goodale, M.A. (2001) The dissociation between perception and action in the Ebbinghaus illusion: non-illusory effects of pictorial cues on grasp, Current Biology, 11, 177-181.

Ingle, D.J. (1973) Two visual systems in the frog, Science, 181, 1053-5.

Jacob, P. & Jeannerod, M. (in press) Ways of seeing, Oxford: Oxford University Press.

Jeannerod, M. (1986) The formation of finger grip during prehension. a cortically mediated visuomotor pattern, Behavioural Brain Research, 19, 99-116.

Jeannerod, M. (1994) The representing brain. Neural correlates of motor intention and imagery, Behavioral and Brain Sciences, 17, 187-245.

Jeannerod, M. (1997) The Cognitive Neuroscience of Action, Oxford: Blackwell.

Kanwisher, N. (2001) Neural events and perceptual awareness, in Dehaene, S. & Naccache, L. (eds.) The Cognitive Neuroscience of Consciousness, Cambridge, Mass.: MIT Press.

Leslie, A. M., Xu, F., Tremoulet, P.D. & Scholl, B.J. (1998) Indexing and the object concept: developing 'What' and 'Where' systems, Trends in Cognitive Sciences, 2, 1, 10-18.

Marshall, J & Halligan, P.W. (1994) The yin and yang of visuospatial neglect. A case study, Neuropsychologia, 32, 1037-1057.

McDowell, J. (1994) Mind and the World, Cambridge, Mass.: Harvard University Press.

Millikan R.G. (1996) Pushmi-pullyu representations, in Tomberlin, J. (ed.) Philosophical Perspectives, vol. IX, Atascadero CA: Ridgeview Publishing.

Milner, A.D. & Goodale, M.A. (1995) The Visual Brain in Action, Oxford: Oxford University Press.

Milner, A.D., Paulignan, Y., Dijkerman, H.C., Michel, F. and Jeannerod, M. (1999) A paradoxical improvement of misreaching in optic ataxia: new evidence for two separate neural systems for visual localization, Proceedings of the Royal Society, 266, 2225-9.

Milner, A.D., Harvey, M., Roberts, R.C. & Forster, S.V. (1993) Line bisection errors in visual neglect: misguided action or size distortion, Neuropsychologia, 31, 39-49.

Milner, A.D., Perrett, D.I., Johnston, R.S., Benson, P.J., Jordan, T.R., Heeley, D.W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E. & Davidson, D.L.W. (1991) Perception and action in 'visual form agnosia', Brain, 114, 405-428.

Mountcastle, V.B., Lynch, J.C., Georgopoulos, A. Sakata, H. and Acuna, C. (1975) Posterior parietal association cortex of the monkey: command functions for operations within extra-personal space, Journal of Neurophysiology, 38, 871-908.

Murata, A., Gallese, V., Luppino, G., Kaseda, M. and Sakata, H. (2000) Selectivity for the shape, size, and orientation of objects for grasping in monkey parietal area AIP, Journal of Neurophysiology, 79, 2580-2601.

Needham, A. (2001) Object recognition and object segregation in 4.5-month-old infants, Journal of Experimental Child Psychology, 78, 3-24.

Noë, A. (2002) Is the visual world a grand illusion?, in Noë, A. (ed.) Is the Visual World a Grand Illusion?, Thorverton: Imprint Academic.

O'Regan, J.K. & Noe, A. (2001) A sensorimotor account of vision and visual consciousness, Behavioral and Brain Sciences, 24, 939-1031.

O'Shaughnessy, B. (1980) The Will, Cambridge: Cambridge University Press.

O'Shaughnessy, B. (1992) The diversity and unity of action and perception, in Crane, T. (ed.) The Contents of Experience, Cambridge: Cambridge University Press, 216-66.

Peacocke, C. (1992) A Study of Concepts, Cambridge, Mass.: MIT Press.

Peacocke, C. (1998) Nonconceptual content defended, Philosophy and Phenomenological Research, 58, 381-88.

Peacocke, C. (2001) Does perception have nonconceptual content, The Journal of Philosophy, 98, 239-64.

Piattelli-Palmarini, M. (ed.)(1980) Language and Learning: the Debate Between Jean Piaget and Noam Chomsky, Cambridge, Mass.: Harvard University Press.

Pisella, L, Gréa, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., Boisson, D. & Rossetti, Y. (2000) An 'automatic pilot' for the hand in human posterior parietal cortex: toward a reinterpretation of optic ataxia, Nature Neuroscience, 3, 7, 729-36.

Pylyshyn, Z. (2000a) Situating vision in the world, Trends in Cognitive Science, 4, 5, 197-207.

Pylyshyn, Z. (2000b) Visual indexes, preconceptual objects and situated vision, Cognition, 80, 127-58.

Rossetti, Y. & Pisella, L. (2002) Several 'vision for action' systems: a guide to dissociating and integrating dorsal and ventral functions (tutorial), in Prinz, W. & Hommel, B. (eds) Common Mechanisms in Perception and Action, Oxford University Press.

Schneider, G.E. (1967) Contrasting visuomotor functions of tectum and cortex in the golden hamster, Psychologische Forschung, 31, 52-62.

Searle, J.R. (1983) Intentionality. An essay in the philosophy of mind. Cambridge University press, Cambridge.

Tanaka, K. (1993) Neuronal mechanisms of object recognition, Science, 262, 685-688.

Teuber, H.L. (1960) Perception, in Field, J., Magoun, H.W. and Hall, V.E. (eds.) Handbook of Physiology, Section I, Neurophysiology, American Physiological Society, Washington, pp. 89-121.

Tye, M. (1995) Ten Problems of Consciousness, Cambridge, Mass.: MIT Press.

Ungerleider, L. & Mishkin, M. (1982) Two cortical visual systems, in Ingle, D.J. Goodale, M.A. and Mansfield, R.J.W. (eds) Analysis of visual behavior, MIT Press, Cambridge, pp. 549-586.

Wynn, K. (1992a) Addition and subtraction by human infants, Nature, 358, 749-50.

Wynn, K. (1992b) Evidence against empiricist accounts of the origins of numerical knowledge, Mind and Language, 7, 4, 315-32.

Xu, F. & Carey, S. (1996) Infants' metaphysics: the case of numerical identity, Cognitive Psychology, 30, 111-53.

Zeki, S. (1993) A Vision of the Brain, Blackwell, Oxford.

Caption of Figure 1

A view of the anatomical bifurcation between the ventral and the dorsal pathways in the primate visual system originating from the primary visual cortex (areas V1 and V2). The ventral pathway projects towards the inferotemporal cortex (areas TE and TEO). The dorsal pathway projects towards the parietal cortex. Many of the connections are reciprocal. (From Boussaoud et al., 1990.)

Caption of Figure 2

Single cell recorded in inferotemporal area of the brain of a macaque monkey. The neuron responds to the sight of a tiger's head viewed from above (as represented in lower left). The ordered sequence of arrows shows the progressive simplification of the stimulus that keeps triggering the discharge of the neuron. It still responds to a white square flanked by two black rectangles. Further reduction fails to trigger cell response. (From Tanaka, 1993.)

Caption of Figure 3

Single cell recorded in the intraparietal sulcus of the brain of a macaque monkey.

(a) shows different responses of a cell to objects with different geometrical properties in object fixation task. Note maximal response in A and lowest in D.

(b) shows different responses of a cell in a visuomotor task. Note maximal response in A. (From Murata et al., 2000.)

Caption of Figure 4

(a) shows unfolding of hand movement in grasping task in healthy subject.

(b) shows abnormal pattern of opening and closure of fingers in grasping task in optic ataxi patient.

(c) shows lack of both preshaping of finger grip and misreaching in reaching task in optic ataxic patient. (From Jeannerod, 1986.)

Caption of Figure 5

(a) and (b) show the performances of, respectively, an apperceptive patient and a healthy subject, in a perceptual task in which the subject is asked to match the orientation of a slot by orienting her wrist over successive trials.

(c) and (d) respectively show the performances of, respectively, an apperceptive patient and a healthy subject, in a visuomotor task in which the subject is asked to insert a hand-held card through a slot at different orientations. (From Goodale, 1995.)

Caption of Figure 6

Subjects are asked to move their index finger from the closer to the more distant vertex of a Müller-Lyer segment. (Adapted from Gentilucci et al., 1996).

Caption of Figure 7

The Titchener circles (or Ebbinghaus) illusion

In (a), the circle surrounded by an annulus of smaller circles looks larger than a circle of equal diameter surrounded by an annulus of larger circles.

In (b), the circle surrounded by the annulus of smaller circles is smaller, but looks equal, to the circle surrounded by an annulus of larger circles.

Caption of Figure 8

(a) shows maximum grip aperture in a visuomotor task of grasping a disk in a Titchener circles condition.

(b) shows manual estimation of the size of a disk in same condition. Note the sensors attached to thumb and index finger. (From Haffenden and Goodale, 1998.)

Caption of Figure 9

Leftmost square shows small distance between central disk and annulus of smaller circles in standard Titchener circles illusion. Rightmost square shows large distance between central

disk and annulus of larger circles in standard Titchener circles illusion. Central square shows large distance between central disk and annulus of smaller circles in non-standard Titchener circles illusion. (From Haffenden, Schiff and Goodale, 2001.)

Caption of Figure 10

White bars show manual estimation for each of the three conditions displayed separately: traditional small, traditional large and adjusted small (see Figure 9). Black bars show maximal grip aperture in grasping task in each of three condition. The experiment shows dissociation between manual estimation and grasping task: manual estimation of adjusted small is very much like manual estimation of traditional small. But grasping of traditional small is very much like grasping of traditional large. (From Haffenden, Schiff and Goodale, 2001.)

Caption of Figure 11

(A) shows high level of extinction by a neglect patient for presentations of stimulus that fails to elicit a Kanizsa square illusion.

(B) shows much lower extinction by the same neglect patient for presentations of stimulus that do elicit a Kanizsa square illusion. (From Driver and Vuilleumier, 2001.)

(C) shows that as the patient's fixation point was in the middle of the card ("bilateral events") in either A or B, she was visually aware of the Kanizsa square (low rate of extinction: black rectangle on the right); but she was not visually aware of the four black circles with white quarter segments (high rate of extinction: large black rectangle on the left).