

Visual cognition: a new look at the two-visual systems model

Marc Jeannerod, Pierre Jacob

► **To cite this version:**

| Marc Jeannerod, Pierre Jacob. Visual cognition: a new look at the two-visual systems model. 2004.
| <ijn_00000569>

HAL Id: ijn_00000569

https://jeannicod.ccsd.cnrs.fr/ijn_00000569

Submitted on 27 Jan 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Research report

Visual cognition: a new look at the two-visual systems model

M. Jeannerod^{a,*}, P. Jacob^b^a *Institut des Sciences Cognitives, Lyon, France*^b *Institut Jean Nicod, Paris, France*

Abstract

In this paper, we argue that no valid comparison between visual representations can arise unless provision is made for three critical properties: their direction of fit, their direction of causation and the level of their conceptual content. The conceptual content in turn is a function of the level of processing. Representations arising from earlier stages of processing of visual input have very little or no conceptual content. Higher order representations get their conceptual content from the connections between visual cognition and other parts of the human cognitive system. The two other critical properties of visual representations are their mind/world direction of fit and their mind/world direction of causation. The output of the semantic processing of visual input has a full mind-to-world direction of fit and a full world-to-mind direction of causation: it visually registers the way the world is and is caused by what it represents. The output of the pragmatic processing yields information for the benefit of intentions, which clearly have a world-to-mind direction of fit and a mind-to-world direction of causation. An intention is both the representation of a goal and a cause of the transformation of a goal into a fact. These properties segregate representations specialized for perception from those specialized for action. Perception implies comparison between simultaneously represented and analyzed objects: hence, object perception presupposes the representation of spatial relationships among objects in a coordinate system independent from the perceiver. Spatial relationships carry cues for attributing meaning to an object, so that their processing is actually part of semantic processing of visual information. These considerations lead to a re-evaluation of the role of the two classical pathways of the human visual system: the ventral and the dorsal cortical pathways. The parietal lobe, which has been identified with the dorsal pathway, cannot be considered as a unitary entity with a single function. The superior parietal lobule carries visuomotor processing, a non-lateralized process. The right inferior parietal lobule contributes to the perception of spatial relationships, a process with a mind-to-world direction of fit and a world-to-mind direction of causation. Finally, the left inferior parietal lobule contributes to still another type of representation, related to visually goal-directed action, i.e., with both a world-to-mind direction of fit and a mind-to-world direction of causation.

© 2004 Published by Elsevier Ltd.

Keywords: Visual cognition; Two-visual systems hypothesis; Representations; Goal-directed movements; Mental imagery

1. Introduction: the emergence of the two-visual systems hypothesis

Although seeing is commonly experienced as a unitary activity, the scientific understanding of human vision resists such a simple view. Both psychologists and neuroscientists consider that the processing of visual information is distributed across several different routes which eventually reach different functional outcomes, and that these processing routes can be mapped onto well-identified anatomical subdivisions of the visual system. This general idea finds support

in the anatomical organization of the visual system in all the vertebrate species that have been studied over the last hundred years, including frogs, fishes, cats, rats, bats, tree-shrews or monkeys, where the retina projects onto many different cortical and subcortical relays. Indeed, the early versions of the two-visual systems hypothesis were first entertained by neurophysiologists working on the visual systems of non-human animals. In amphibians, for example, 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 (1969) that a hamster with a lesioned superior colliculus could discriminate vertical from horizon-

* Corresponding author. Tel.: +33 3791 1212; fax: +33 3791 1210.
E-mail address: jeannerod@isc.cnrs.fr (M. Jeannerod).

tal 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 vision controlled by peripheral retinal information, based on subcortical structures, and vision based on cortical structures, respectively.

The first major step was taken by Ungerleider and Mishkin (1982), who located the two-visual systems within the primate visual cortex. They 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 containing a reward—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 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 concluded that both the ventral stream (which they called the 'object-channel') and the dorsal stream (which they called the 'space-channel') were specialized in perceiving different aspects of the visual world. Indeed, their landmark task tested the animal's ability to *perceive* spatial relations, not to act on a target.

The second major step was taken by 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. The visuomotor transformation is the automatic conversion of visual information into hand commands for reaching and grasping objects—a topic whose study was pioneered in the monkey by Mountcastle and collaborators in the mid-1970s. In Milner and Goodale's 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: a form of visual agnosia resulting from lesions in the inferotemporal area and optic ataxia resulting from lesions in the posterior parietal cortex. Visual agnostic patients are deeply impaired in the visual recognition of the color, size, shape and orientation of objects. But they can reach and grasp objects they cannot visually recognize. Conversely, optic ataxic patients fail to reach and grasp objects whose shapes, sizes and orientations they can visually recognize.

Many relevant psychophysical experiments have confirmed the view that one and the same visual stimulus can

be processed in two different ways according to the task by healthy human subjects. Consider, for example, the common illusion of visual 'induced motion': a small stationary visual target is presented on a screen against a large background of dots constantly moving in one direction. A subject located in front of the screen will report that the small target appears to be moving in the direction opposite to the background. If, however, the subject is instructed to point at the small target (with his unseen hand in order to avoid visual guidance of the pointing movement), he will accurately reach the actual position of the target. As this example (Bridgeman, Kirsch, & Sperling, 1981) shows, visual perception and visually guided action can be dissociated by carefully designed experiments in normal subjects.

Thus, by the mid-1990s, 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 visual perception: the two cortico-cortical pathways in the primate visual brain underlie perceptual awareness. By contrast, according to Milner and Goodale's (1995) model, perceptual awareness is not the exclusive (or the main) function of vision in primates. Cortico-cortical pathways in the primate and the human brains are not limited to visual perception.

We do accept Milner and Goodale's (1995) basic bifurcation between vision-for-perception and vision-for-action, which we call, respectively, the "semantic" and the "pragmatic" processing of visual information (see, e.g., Jeannerod, 1997). We shall argue that Milner and Goodale's model of the two-visual systems hypothesis seriously underestimates the complexity of the representations of actions produced by the pragmatic processing of visual information. No doubt, one of the functions of the dorsal stream is to enable the visuomotor transformation. But, as we shall argue in this paper, the human parietal lobe has two other major functions: one is to allow the perception of spatial relations among objects; the other is to store complex representations of actions (such as schemas for the use of cultural tools). On the perceptual (or semantic) side, it would be absurd to reduce the scope of human vision to the perception of objects that one can manipulate with one's hand. Humans can of course visually perceive a great variety of other things such as clouds, flames, shadows, holes and many others. In particular, humans can also visually perceive actions performed by conspecifics. Similarly, the visual control of human actions should not be restricted to the visuomotor transformation, i.e., to reaching and grasping objects. Humans can plan, execute and visually control far more complex actions.

2. The architecture of visual cognition

The basic insight of the two-visual systems hypothesis is that the goal of visual processing is two-fold: on one hand, hu-

man visual perception is a rich source of knowledge about the world; on the other hand, human vision contributes to visually guided actions on the world. The computational requirements, respectively, of perception and the control of object-oriented actions on the human visual system are clearly different. Perception itself fulfills two complementary functions: selection and recognition. The selection phase consists in both segregating a complex visual array into several separable objects and in attributing to each separate object its own set of appropriate visual attributes (this is the so-called “binding” problem). Usually, the color and texture of an object will be highly relevant to its perceptual selection from a set of neighboring objects. Segregation and binding require that the relative spatial locations of different objects in a visual array be coded by the perceptual system. Since perceptual recognition of an object must be achieved from many different spatial perspectives on many different occasions, it requires encoding of visual information about an object’s enduring properties. In other words, perceptual recognition of an object demands that visual information about a perceived object matches conceptual information and knowledge about it stored in long-term memory.

Once an object has been perceptually selected from a set of competitors, the visual control and monitoring of the action of prehension can take over. Whereas the color and texture of an object are relevant to its perceptual selection, they are not relevant to grasping it. What is relevant to the visual guidance of grasping an object are its absolute shape, size and orientation together with its position relative to the agent’s body.

Two of the most fundamental dimensions along which visual percepts differ from visuomotor representations are what, following Anscombe (1957) and Searle (1983), philosophers of mind and perception call, respectively, their direction of fit and their direction of causation (for a full account of this distinction, see Jacob & Jeannerod, 2003). Beliefs and visual percepts are descriptive representations. They have a mind-to-world direction of fit: their job is to represent facts or actual states of affairs. If what a belief or a percept represents fits a mind-independent fact in the world, then the belief or the percept is veridical; otherwise, they are not. By contrast, intentions and desires are prescriptive representations. They have a world-to-mind direction of fit: their job is to represent goals, i.e., possible or impossible (i.e., non-actual) states of affairs. If what obtains in the world fits what the intention or the desire represents, then the intention or the desire is fulfilled; otherwise, they are not.

In addition, percepts and intentions have an opposite mind–world direction of causation. If you perceive a basket full of apples, pears, lemons and oranges, your visual percept is caused by the state of affairs that it represents. Unless there was a basket full of apples, pears, lemons and oranges, you might hallucinate one, but you could not perceive it. Whereas visual percepts are caused by what they represent, intentions cause the state of affairs which they represent. When you intend to pick up an apple from the perceived basket full of fruits

in front of you, you represent a goal for action, i.e., a possible non-actual state of affairs that consists of your holding a particular apple in your hand. When you form the intention to grasp the apple, the perceived fact is that the target of your action (i.e., the apple) is in the basket, not in your hand. Your intention causes your action which in turn causes a new fact to obtain, i.e., the apple’s being in your hand.

Clearly, visual percepts and beliefs have the same mind-to-world direction of fit. The formation of beliefs about the world is a step towards the acquisition of knowledge about the world. While beliefs have a conceptual content, visual percept has a rich pictorial non-conceptual content. The function of visual percepts is to provide visual information relevant to the formation of beliefs, and thus of knowledge about the visual properties of the world. The philosopher Millikan (1996) has argued that there exists, in the human mind (and in the minds of other animals), a class of Janus-like mental representations, which she calls “pushmi-pullyu” representations—after the Pushmi-Pullyu, an imaginary two-headed animal in Dr. Doolittle’s stories. Visuomotor representations are such representations with a hybrid direction of fit in virtue of which they provide motor intentions, not beliefs, with visual information about affordances for action. Because they represent only immediate affordances for action, the non-conceptual content of visuomotor representations is not as rich as the non-conceptual content of visual percepts.

The contrast between the direction of fit, respectively, of visual percepts and visuomotor representations is confirmed by the neuropsychological double dissociation already mentioned between the perceptual impairment of visual agnostic patients and the visuomotor impairment of optic ataxic patients. The visual form agnostic patient DF cannot form visual percepts, but she can still form visuomotor representations of targets of hand actions (Goodale, Milner, Jakobson, & Carey, 1991). Conversely, optic ataxic patients cannot form visuomotor representations of targets of hand action, but they can form visual percepts (Jeannerod, 1986).

3. Levels of semantic processing of visual information

What we call “semantic” processing of visual information is the process whereby visual inputs are transformed into perceptual representations with a mind-to-world direction of fit and whose pictorial non-conceptual contents must ultimately match the conceptual contents of beliefs. As we pointed out above, the goal of semantic processing of visual inputs is recognition of objects which involves segregation of a scene into separable objects and binding to each object of its appropriate visual attributes. Only representations with a fairly abstract conceptual content can be used in thinking and reasoning about objects. In order to match the conceptual contents of general thoughts and beliefs stored in memory, much of the detailed pictorial content of visual percepts must be selectively eliminated.

At the lowest level, perception can be, as philosophers of perception (e.g., Dretske, 1969, 1978) call it, non-epistemic. Suppose you are driving very fast. You see something lying on the road ahead of you. You cannot identify it because you are moving too fast, but you nonetheless skillfully avoid hitting it. Unless you saw it, you would have hit it. So you did see it, but you could not see what it was. Your visual perception of the object on the road was non-epistemic. Epistemic visual perception involves further processing of an object giving rise to some identification: one sees epistemically, not just an object, but an object as instantiating some category or other. Alternatively, one sees epistemically the fact that an object falls under some concept or other. Seeing that the car in the street is moving at a slow speed, for example, is forming a visual percept of an object with a global contour and several parts of different shapes, colors and textures that move together relative to other surrounding objects. For the purpose of considering the functional advantages or disadvantages of, e.g., a Ford Mustang with cars of other brands, however, one must switch from the detailed pictorial representation of a visual percept to more general knowledge about cars sustained by representations with a more abstract conceptual content (e.g., the concept of a Ford Mustang).

Similarly, one cannot visually perceive a mug of beer as being to the left of a bottle of wine without representing, e.g., the particular shades of colors and the levels of the liquids contained in both the mug and the bottle, and the particular shapes of the mug and the bottle. Now, a mug of beer can only be seen to be to the left of a bottle of wine from some spatial point of view, e.g., from the point of view of someone facing the window, not from the point of view of someone with her back onto the window. From the latter point of view, one will see the same mug of beer to the right, not to the left, of the same bottle of wine. Now, one can think about, but one cannot currently see, the point of view by means of which one is currently seeing a mug of beer as being to the left of a bottle of wine. One may see at t the point of view one occupied at $t - 1$, by occupying at t a point of view different from the one is currently seeing at t . So in order to form the thought that the mug of beer is to the left of the bottle of wine from the point of view of someone facing the window, not from the point of view of someone with his back onto the window, one must ascend from a visual percept with a rich pictorial non-conceptual content to a thought with a more abstract conceptual content. One must abstract away from the pictorial content of the visual percept representing, e.g., the colors and levels of the two liquids and the shapes of the mug and the bottle, in order to form a conceptual representation of the spatial relation “ x is to the left of y from z 's point of view”.

Much of our knowledge about the mechanisms underlying visual recognition and identification of objects comes from the observation of patients with brain lesions. Clinical observation has provided information which could have hardly been obtained from studies on subjects with an intact brain. Specifically, clinical observation gives firm support to the

idea of the existence of levels of perceptual processing of visual inputs. The higher levels of processing correspond to the common experience of seeing, which can be easily studied in normal subjects: this is the goal of cognitive perceptual psychology. By contrast, the lower levels, which correspond to covert stages of processing, become only apparent in patients where a lesion has impaired the higher levels.

Indeed, visual identification disorders resulting from lesions of the equivalent of the ventral pathway in the human visual system correspond to impairments of visual information processing at different levels. When they are bilateral, posterior lesions affecting the lateral occipital region destroy visual representations resulting from an early stage of perceptual processing (what philosophers call non-epistemic visual perception). Such representations have a rich informational content and little or no conceptual content. The resulting effect (visual form agnosia) is that patients with such lesions cannot form simple percepts from the visual array: they cannot recognize the simple shapes, orientation and color of visual stimuli. As a consequence of this impairment, the more cognitive visual representations (those with a conceptual content) are ‘deafferented’ from visual input and cannot achieve their task of object recognition.

Several observations, however, suggest that these higher representations may still be functional. Servos and Goodale (1995), for example, found that the visual form agnostic patient DF had retained the ability to form visual mental images of objects: although she could not recognize visually presented objects and could not draw copies of seen objects, she could draw copies of objects from memory—which she then could hardly recognize. Patients with even more posterior occipital lesions including lesions of the primary visual areas in the calcarine sulcus, who present the typical picture of cortical blindness, spontaneously report vivid visual images (and sometimes even deny being blind, see Goldenberg, Müllbacher, & Nowak, 1995).

More anterior lesions (e.g., bilateral or predominantly left-sided lesions of the inferotemporal cortex) destroy more cognitive representations with conceptual content, those that give access to the meaning of the percepts, and allow processes like comparison and categorization. The resulting effect is “associative agnosia”, a condition in which patients often retain the ability to identify simple shapes and are even able to copy line drawings of objects that they cannot recognize. Whereas presemantic recognition of objects is preserved in these patients (they are able to form visual percepts), full semantic identification seems to be lost. Not surprisingly, these patients are usually unable to perform mental visual imagery tasks (although there are several well-documented exceptions to this rule, see Berhmann, Winocur, & Moscovitch, 1992). The loss of the ability to mentally image visual objects (e.g., faces) is congruent with the findings obtained with neuroimaging techniques from normal subjects during mental visual imagery tasks. Typically, these tasks activate brain areas at the occipitotemporal junction as well as in the inferotemporal cortex (see Farah, 1995): the activated ar-

384 eas superimpose with those activated during recognition and
 385 matching of seen objects (e.g., Ungerleider & Haxby, 1994).
 386 These results, together with the effects of anterior lesions of
 387 the ventral pathway, thus suggest that the inferotemporal cor-
 388 tex might be a critical site for semantic processing, including
 389 long-term declarative memory, of visual objects. The net-
 390 work for the generation of mental visual imagery of objects,
 391 as described with the use of neuroimaging techniques in nor-
 392 mal subjects, also includes more posterior occipital areas,
 393 including areas in the calcarine sulcus (Kosslyn et al., 1993;
 394 Kosslyn, DiGirolamo, Thompson, & Alpert, 1998). Although
 395 at first sight this result seems in conflict with preservation of
 396 the ability to evoke such images in patients with occipital
 397 lesions, it is possible that a more complete analysis of visual
 398 mental imagery in agnosic patients with posterior lesions will
 399 reveal subtle impairments with respect to the normal process.

400 The feed-forward mode of information processing is often
 401 considered as the main (if not the only) constituent of visual
 402 cognition. Indeed, in his own definition of visual cognition,
 403 Pinker (1985) states that it can be conveniently divided into
 404 two serially organized steps which indeed fulfill our crite-
 405 ria for a world-to-mind direction of fit. Pinker's first step "is
 406 the representation of information concerning the visual world
 407 currently before a person [. . .] the process that allows us to
 408 determine on the basis of retinal input that particular shapes,
 409 configurations of shapes, objects scenes and their properties
 410 are before us". The second step "is the process of remember-
 411 ing or reasoning about shapes or objects that are not currently
 412 before us but must be retrieved from memory or constructed
 413 from a description" (pp. 2–3). Whereas Pinker's first stage
 414 is consistent with the world-to-mind direction of causation
 415 we ascribed to visual percepts, Pinker's second stage is more
 416 like beliefs, which may well fail the world-to-mind direction
 417 of causation.

418 4. Levels of pragmatic processing of visual 419 information

420 As we suggested in Sections 1 and 2, not all visual rep-
 421 resentations are percepts with a world-to-mind direction of
 422 fit. They do not all result from semantic processing. There
 423 are also visuomotor representations that result from the prag-
 424 matic processing of visual inputs. As we argued above, vi-
 425 suomotor representations have a hybrid direction of fit that
 426 makes them suitable for providing motor intentions with vi-
 427 sual information about targets of action. Unlike percepts and
 428 beliefs, intentions have a mind-to-world direction of fit and
 429 unlike percepts, they have a mind-to-world direction of cau-
 430 sation: they cause bodily movements that turn a possible into
 431 an actual state of affairs.

432 For the purpose of introducing the notion of pragmatic
 433 processing, we focused on visuomotor representations that
 434 are involved in the visuomotor transformation, i.e., in the vi-
 435 sual control of reaching and grasping objects. What is crucial
 436 to the content of such visuomotor representations is that they

437 represent the geometrical properties of objects relevant for
 438 grasping and that they code the spatial position of the tar-
 439 get in egocentric coordinates, i.e., in a frame of reference
 440 centered on the agent's body. The visuomotor transforma-
 441 tion is but the lowest level of pragmatic processing of visual
 442 inputs. Visuomotor representations of targets of prehension
 443 are representations with little or no conceptual content at all.
 444 The scope of pragmatic processing, however, is not limited
 445 to the visuomotor transformation, since pragmatic process-
 446 ing is involved in conceptually more complex operations like
 447 evaluating the feasibility of an action, anticipating its conse-
 448 quences, planning further steps and learning the skilled use
 449 of tools by observation. Such representations include concep-
 450 tual information about previous experience (hence memory),
 451 about the context in which the action has to be performed
 452 (e.g., danger, competition), up to its moral implications (if
 453 any).

454 The study of visuomotor behavior already reveals that even
 455 simple goal-directed movements are likely to be represented
 456 by the agent prior to their execution. Consider, for exam-
 457 ple, the action of grasping with the right hand a horizontally
 458 placed rod. Prior to his movement, the subject receives an
 459 instruction about what to do after the rod has been grasped:
 460 the instruction is (according to trials) either to place the right
 461 end or the left end of the rod on a stool. These instructions
 462 generate a highly consistent behavior. When the instruction
 463 is to place the right end of the rod on the stool, the subject
 464 invariably uses an overhand grip; conversely, for placing the
 465 left end of the rod on the stool, the subject uses an under-
 466 hand grip (Rosenbaum & Jorgensen, 1992). This process of
 467 grip selection (a typical example of visually based decision)
 468 suggests that biomechanical constraints generated by grasp-
 469 ing the object and rotating the wrist are encoded within the
 470 representation of the movement: it is easier to rotate the hand
 471 in the pronation direction than in the supination direction
 472 (see Stelmach, Castiello, & Jeannerod, 1994). An alterna-
 473 tive explanation for this behavior would be that the visual
 474 configuration of the rod and the stool simply affords a pronation
 475 movement which is directly executed without building a
 476 representation. This alternative can be ruled out. Indeed, the
 477 same categorical decision is observed in a situation where
 478 the action of moving the rod to the stool is imagined (or
 479 simulated) but not executed (Johnson, 2000). Thus, mentally
 480 simulated hand movements follow the same rules and obey
 481 the same constraints as their overtly executed counterparts.

482 This finding was first reported in Parsons' (1994) hand
 483 matching experiment. In this experiment, a subject is shown
 484 the image of a sample hand in its canonical orientation. An-
 485 other hand (the test hand) is then briefly presented at a dif-
 486 ferent orientation and/or in a different posture. The subject's
 487 task is to tell whether or not the laterality (right or left) of the
 488 test hand matches that of the sample hand. The time taken
 489 by the subject to give the response is found to be a func-
 490 tion of the difference in orientation between the two hands.
 491 So far, this result is in line with the well-known mental ro-
 492 tation phenomenon described in mentally matching two 3D

visual shapes (e.g., Shepard & Metzler, 1971). Unlike a 3D visual shape, however, one cannot rotate one's own hand in any direction: instead, the rotation of one's hand to a given orientation has to follow biomechanically compatible trajectories. This constraint is reflected in the results of the above hand matching experiment, where the response time is also a function of the compatible trajectory of the test hand, as if the subject were actually rotating his own hand. Other experiments of the same vein and using the same methodology of mental chronometry have confirmed that mentally imagined movements follow the same regularities as those which have been described for executed movements, for example, simulated reaching follows Fitts' law (Decety & Jeannerod, 1996; Sirigu et al., 1996).

Thus, visuomotor representations appear to have a relatively direct influence on motor mechanisms, i.e., those mechanisms involved in the execution of the represented movements. This hypothesis is supported by the results of experiments where brain activity is monitored during cognitive tasks such as making visually based decisions, forming motor images or remembering motor events. These experiments (Decety et al., 1994; Parsons et al., 1995; Nilsson et al., 2000; Johnson et al., 2002; Shubotz & von Cramon, 2002) reveal that, in the absence of any movement or muscular activity, brain areas corresponding to motor areas are activated. At the cortical level, primary sensorimotor cortex and dorsal and ventral premotor cortex are activated as well as, subcortically, the lateral cerebellum and basal ganglia. The activated structures partly but consistently overlap those that are activated during actual motor performance of the same actions (Gérardin et al., 2000; see review in Jeannerod & Frak, 1999). Not surprisingly, increased neural activity in motor areas observed during a simulated action (e.g., mental hand rotation) is not observed during mental rotation of visual shapes, which only affects visual areas (Kosslyn et al., 1998).

Visuomotor representations and their close connections to motor execution, however, are only one among the possible classes of visual representations built for acting on the world. Their role is also to feed in more complex representations, more remote from visual input but which include more conceptual content. The contents of representations that result from higher level pragmatic processing include contextual elements drawn from the situation in which the action is taking place, such as the precise function of the objects which are part of this action. Consider, for example, a skilled action using tools. Tools, as well as musical instruments or sport materials, are objects which cannot be characterized merely by their geometrical properties like size, shape or orientation. They have additional properties that cannot be detected unless one knows what the object is for and how to use it; yet, once they are known (by observation, training or verbal instructions), they do supervene upon the pure geometrical properties that are part of the non-conceptual content of more basic visuomotor representations. Thus, the use of tools, the practice of musical instruments or the use of sport materials require the construction of visual representations whose con-

ceptual content results from the pragmatic processing of visual information. For example, being able to use a tool and/or a musical instrument depends upon observing the skilled actions performed by others.

Watching other people act is indeed a source of information about the meaning of their actions and, ultimately, about the contents of their mental states. Observation of an action (e.g., performed with a tool or a musical instrument) first provides clues about the technical aspects of that action, for learning and replicating it. But observing an action and understanding its goal may also provide information about the agent's intentions, desires and motives. The capacity to imitate (which seems so distinctly human) depends on the ability to form visual representations of others' observed actions. Some perceptual representations of object-oriented actions play a crucial role in learning how to use such things as tools or musical instruments. They contribute to understanding the agent's motor intentions. One's representations of others' object-oriented actions share many of the neural correlates of one's own visuomotor decisions or of one's own imagined actions and motor imagery. This fact lends support to the idea that covert action or mental simulation is at work as well in the preparation of one's own object-oriented actions as in the perception and understanding of others' object-oriented actions (Jeannerod, 2001).

One may also watch a conspecific act, not in the context of learning a skilled action, but in a different social and emotional context. One may, for example, either watch the arm and hand movements of a person engaged in a fist-fight for the purpose of learning how to fight or in the social context of witnessing the action of an aggressor inflicting pain on a victim. Such a perceptually based representation of another's action encodes a wealth of visual stimuli endowed with a social, not a motoric (or technical), significance. In this case, the observed "actions", which are directed towards conspecifics, not towards inanimate objects, may include facial expressions, eye movements and fixations, changes in posture, or gestures which are not directed to external objects or goals, but which have a social ostensive or demonstrative role. Whereas perceptually based representations of object-oriented actions contribute to determining and understanding the agent's motor intention, perceptually based representations of actions directed towards conspecifics contribute to determining and understanding the agent's social intention, i.e., the agent's intention to affect a conspecific behavior. We argue elsewhere that, in the human brain, the cortical network associated with the perception of human actions directed towards manipulable objects is distinct from that associated with the perception of human actions directed towards conspecifics (Jacob & Jeannerod, 2003).

As we noted at the end of Section 2, the neuropsychological dissociation between visual agnosia and optic ataxia is confirmation of the distinction between the world-to-mind direction of fit of semantic processing and the hybrid direction of fit of pragmatic processing. On the basis of this double dissociation, Milner and Goodale (1995) hypothesized the fol-

lowing duality between the dorsal and the ventral pathway. In their model, the dorsal pathway underlies the visuomotor transformation, i.e., the crude, fast and automatic transformation of information about visual attributes of objects into motor commands. By contrast, the ventral pathway underlies visual perception, i.e., the conscious identification and recognition of objects. Although this model does capture one of the most obvious divisions of labor between visual pathways, it may not be entirely accurate: its main problem is that it wrongly draws a contrast between two kinds of visual information processing located in the two pathways at different levels of complexity. Both semantic processing, which depends on the activity of brain areas in the ventral stream, and pragmatic processing, which depends on the activity of brain areas in the dorsal stream, give rise to representations at different levels, whose content can be more or less conceptualized. Non-epistemic perception stands to semantic processing as the visuomotor transformation stands to pragmatic processing. No conclusion can be drawn about the differences between the semantic and the pragmatic processing of visual inputs unless the levels of conceptual content of their respective outputs are matched.

By contraposition, a valid comparison can be made between higher level representations in both the semantic and the pragmatic systems of processing. Neuropsychology offers a wealth of clinical observations of patients whose higher level representations for visually goal-directed actions are altered and whose behavior could be compared with the behavior of patients with deficits in semantic processing. The difficulties met by these patients appear in situations where they have to use tools for achieving a task on a visual goal. Yet, their impairment is not limited to motor execution: they also typically fail in tasks like pantomiming an action without holding the tool, imitating an action performed by another agent, judging errors from incorrectly displayed actions or imagining an action (motor imagery) (Clark et al., 1994; Sirigu et al., 1995; Ochipa et al., 1997; Goldenberg, Hartmann, & Schlott, 2003). Such impairments in representing actions do not result from a general difficulty in visual recognition: Sirigu and Duhamel (2001) report the cases of two patients whose visual impairments in visual recognition tasks and in motor representations were dissociated. One patient with a left parietal lesion with ideomotor apraxia was unable to perform motor imagery tasks but had normal scores in visual imagery tasks. Conversely, another patient with agnosia for faces and visual objects had no visual imagery but normal motor imagery. Similarly, Tomasino, Rumiati, and Umiltà (2002) report the case of one patient with ideomotor apraxia with a left parietal lesion, who was unable to perform the motor mental imagery task involving hand rotation, whereas he was still able to mentally rotate other visual stimuli.

The clinical observations quoted above stress the role of the parietal cortex in monitoring motor representations. Patients' impairments are produced by parietal lesions located in the angular and supramarginal gyri (the inferior parietal

lobule), i.e., more anterior and ventral than those which produce a visuomotor impairment like optic ataxia. Furthermore, when the lesion is unilateral, it is more often localized in the left hemisphere, a lesional lateralization which is irrelevant to optic ataxia. Indeed, apraxic patients with a lesion of the left inferior parietal lobule have no basic visuomotor impairment: they can correctly reach and grasp objects. Rather, they are impaired in the recognition of tools and in the recognition of actions involving the use of tools. They cannot pantomime actions involving the use of an imaginary tool, nor can they recognize pantomimes executed by others. According to Glover's (in press) recent model, while the superior parietal lobule would be mainly involved in the on-line automatic control of basic visually guided actions towards objects, the left inferior parietal lobule would be involved in the higher level intentional planning of more complex actions involving the retrieval of complex representations thought to be stored precisely in that region. This role of the parietal cortex in action planning becomes even more obvious in the representation of non-executed actions, e.g., in imagined actions or in observing actions performed by another agent. Neuroimaging experiments in normal subjects, some of which have been already mentioned earlier in this paper, show that, besides activating motor areas in common with execution, motor representation tasks consistently activate areas in the posterior parietal lobe (Decety et al., 1994; Grafton, Arbib, Fadiga, & Rizzolatti, 1996). In their recent study, Johnson et al. (2002) made a distinction between two mental tasks involving non-executed hand action: grip selection—a simple implicit visuomotor representation—and cued motor preparation, a process which involves attending to one hand and planning a movement with that hand. Grip selection primarily activated a dorsal area of the contralateral parietal lobe, whereas motor preparation activated parietal areas within the left hemisphere. These results are consistent with the above clinical observations, which dissociate visuomotor impairments from impairments in higher level motor representations involved in planning.

The separation, both anatomical and functional, between low-level representations for visuomotor transformation and higher level representations for planning suggests that the function of the occipitoparietal, dorsal, pathway should be refined. The function of the occipitoparietal pathway as described in the monkey, which reaches parietal areas within the intraparietal sulcus and which is connected to premotor areas, is indeed the achievement of the visuomotor transformation. The role of information processing in this pathway is to prepare biomechanically compatible limb trajectories, to compute the speed of the limb movements towards the target, and to adjust the size of the grip and the number of fingers involved for grasping it. These operations are likely to be largely automatic, for the sake of speed and accuracy, although they may be influenced by top-down processing for adapting the movement to the current situation. These points are illustrated by the behavior of patient AT. This patient presented the typical symptoms of optic ataxia exemplified

717 by difficulties during reach and grasp movements: targets
718 presented in her peripheral visual field were misreached, the
719 grip size no longer correlated with object size, the orientation
720 of the opposition axis during grasping no longer correlated
721 with object's orientation (Jeannerod, Decety, & Michel, 1994;
722 Milner, Paulignan, Dijkerman, Michel, & Jeannerod, 1999).
723 However, when presented with familiar objects instead of
724 neutral targets, AT's grasping performance improved signifi-
725 cantly. This effect was likely to be due to a top-down control
726 of the visuomotor transformation, originating from higher
727 order, still intact, representations.

728 By contrast, higher order motor representations with more
729 conceptual content appear to be independent from (though
730 connected with) visuomotor representations. The fact that
731 they are preferentially affected by left-sided lesions indicates
732 that they pertain to a distinct system, concentrated in the
733 inferior parietal lobule. As already mentioned, neuroimag-
734 ing experiments reveal that the regions of the supramarginal
735 gyrus and of the angular gyrus in the inferior parietal lob-
736 ule are activated during tasks involving cued motor plan-
737 ning (e.g., hand selection), motor preparation or mental mo-
738 tor imagery. More recent investigations also found a strong
739 activation of the inferior parietal lobule in tasks involving
740 recognition of one's own actions as opposed to actions per-
741 formed by another self (Ruby & Decety, 2001; Farrer et al.,
742 2003).

743 5. The quasi-conceptual content of the 744 representation of spatial relations

745 To visually represent an object is to represent it in space.
746 One may think about objects that are or that are not in space
747 (e.g., numbers or characters of fiction). But one cannot see
748 objects that are not in space. There are at least three different
749 ways in which one can represent (perhaps non-consciously)
750 the spatial position of an object. All visual processing starts
751 with visual information reaching the retina. So the spatial po-
752 sition of an object is first of all represented in retino-centered
753 coordinates. For the purpose of reaching and grasping an ob-
754 ject, however, the agent must form a representation of its
755 spatial position in egocentric coordinates (i.e., centered on
756 the axis of his body). In a perceptual task, the spatial posi-
757 tion of an object relative to some other neighboring object is
758 represented in allocentric coordinates (i.e., centered on some
759 neighboring object). Thus, part of the duality between the se-
760 mantic processing and the pragmatic processing of a visually
761 presented object derives from the fact that the spatial position
762 of an object can either be coded in allocentric or in egocentric
763 coordinates. Representing the spatial position of an object in
764 egocentric coordinates is required for reaching and grasping
765 it. As we shall presently argue, representing the spatial posi-
766 tion of an object in allocentric coordinates (thus representing
767 its spatial relation to at least one other object present in the
768 visual array) is required for full perceptual awareness of the
769 object's other visual attributes.

770 Consider first experiments by Pylyshyn (2000a, 2000b)
771 about so-called 'multiple object-tracking' (MOT) in normal
772 human adults. First, subjects are shown eight identical circles
773 at rest on a screen, four of which flicker briefly. Secondly, sub-
774 jects see the eight circles move randomly on the screen for
775 about 10 s. Thirdly, they are asked to keep track of the four
776 circles that initially flickered. Normal human adults can keep
777 track of four-to-five such distinct objects (or proto-objects).
778 Now comes the important observation: Pylyshyn reports that
779 subjects fail to notice changes in the colors and shapes of the
780 proto-objects that they tracked by their relative locations. We
781 call 'proto-objects' the circles whose relative motions and
782 positions normal adult subjects manage to track, precisely
783 because such visual features as their shapes and colors are
784 immaterial to their identity in the course of the task. The per-
785 ceptual ability to visually represent the relative motions and
786 locations of proto-objects is impervious to changes of colors
787 and shapes. This ability clearly belongs to semantic process-
788 ing, not to pragmatic processing. Engaging in a MOT exper-
789 iment though seems almost like a cognitive task of spatial
790 reasoning with little or none of the typical phenomenology
791 of visual perception.

792 Neuropsychological studies of patients show that lesions
793 in the dorsal pathway also frequently produce visuospatial
794 impairments. Patients with lesions affecting the posterior
795 parietal areas, usually in the right hemisphere, exhibit spa-
796 tial disorientation: typically, these patients fail to determine
797 the relative locations of objects: they are unable to describe a
798 spatial trajectory and they cannot orient on a map. Lesions lo-
799 cated in the right inferior parietal lobe typically produce uni-
800 lateral spatial neglect. Unlike lesions in the superior parietal
801 lobe, which produce optic ataxia and which can be on either
802 side, lesions responsible for unilateral spatial neglect are gen-
803 erally located in the right hemisphere. Patients with unilateral
804 spatial neglect are not perceptually aware of objects visually
805 presented in their contralesional (i.e., left) hemisphere. For
806 example, when asked to mark line segments at different ori-
807 entations, a neglect patient will systematically fail to mark
808 the segments lying in her contralesional hemifield. If asked
809 to bisect a horizontal line, she will exhibit a strong ipsile-
810 sional bias revealing neglect of the part of the line falling
811 within her neglected hemisphere.

812 Unlike blindsight patients whose primary visual cortex
813 has been damaged, and to a lesser extent unlike visual form
814 agnostic patients whose ventral stream has been impaired,
815 neglect patients lack perceptual awareness on their affected
816 side in spite of the fact that the visual pathway for pro-
817 cessing the neglected visual information remains intact. In-
818 deed, there is considerable evidence for covert processing of
819 the neglected stimuli. For example, Marshall and Halligan
820 (1994) showed neglect patient PS drawings of two houses
821 located on top of each other, one of which displayed brightly
822 colored flames on its left side. When asked to make an
823 explicit comparison between the two houses, the patient
824 could report no difference. When asked, however, which
825 of the two houses she would rather live in, the patient

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

830 These disorders are clearly of a cognitive nature and corre-
831 spond to a failure to build representations of spatial relation-
832 ships between visual objects: one demonstration of this point
833 is provided by the effects of posterior parietal lesions on a
834 special kind of visual imagery (which we tentatively call spa-
835 tial imagery), first described by [Bisiach and Luzzatti \(1978\)](#).
836 These authors reported the case of a patient with left visu-
837 ospatial neglect following a lesion of the right hemisphere,
838 including the parietal lobe. When instructed to build a visual
839 image of familiar surroundings and to describe the content
840 of the image, the patient failed to describe objects located on
841 the left side of his visual image. In other words, the patient
842 seemed to have lost his topographical memory for that lim-
843 ited area of extrinsic space that was visually neglected in his
844 spatial behavior.

845 This observation stresses the fact that what is usually
846 called mental visual imagery should be divided into visual
847 imagery of objects and visual spatial imagery. Whereas the
848 former is involved in representing the visual attributes of ob-
849 jects (such as their color, texture, shape, contour and size) in
850 the absence of retinal inputs, the latter is involved in repre-
851 senting the spatial positions and relations of what we called
852 'proto-objects' in the absence of retinal inputs. There is a dou-
853 ble dissociation between impairments in the visual perception
854 and recognition of objects (produced by inferotemporal les-
855 sions) and impairments in the representation of the spatial
856 relationships between objects (produced by posterior pari-
857 etal lesions). [Levine, Warach, and Farah \(1985\)](#) and [Farah,
858 Hammond, Levine, and Calvanio \(1988\)](#) report similar dis-
859 sociations between the visual imagery of objects and visual
860 spatial imagery. Patients with impaired visual object recogni-
861 tion are also impaired for visual object imagery, whereas their
862 ability for spatial imagery may be preserved. Conversely, pa-
863 tients with spatial disorientation are impaired in spatial im-
864 agery, but not in visual object imagery. Consider, for example,
865 the associative agnosic patient LH described by [Farah et al.](#)
866 Following a bilateral lesion of the occipitotemporal junction
867 and of the inferotemporal cortex, this patient was deeply im-
868 paired in visual recognition for faces, animals, plants, food
869 and many common objects. He was tested in a variety of tasks
870 requiring visual imagery. He was asked about some of the
871 characteristics of well-known objects that are rarely encoded
872 in verbal memory and that require access to iconic memory
873 such as: What is the color of a football?, Do beavers have
874 long tails?, etc. LH was deficient in all these tasks. His deficit
875 in visual object imagery, however, stood in contrast with his
876 preserved ability for spatial imagery. Thus, LH was able to
877 perform mental spatial tasks such as mental rotation of 3D
878 letters or mental scanning.

879 The fact that visual spatial imagery was preserved in this
880 patient has two important implications. First, it is congru-
881 ent with the sparing of his parietal lobes, the role of which

882 in spatial imagery was clearly demonstrated in normal sub-
883 jects with neuroimaging (e.g., [Kosslyn et al., 1998](#)). Neu-
884 roimaging studies involving perceptual and visuospatial tasks
885 (e.g., judgment of relative spatial location and orientation
886 of two or more objects) also consistently show activation
887 of relatively posterior and ventral parietal areas on the right
888 side, in the fundus of the intraparietal sulcus ([Haxby et al.,
889 1994](#); [Faillenot, Decety, & Jeannerod, 1999](#)) as well as in
890 the area of the angular gyrus in the inferior parietal lob-
891 ule ([Köhler, Kapur, Moscovitch, Winocur, & Houle, 1995](#)).
892 Second, this clinical fact demonstrates that visual process-
893 ing in the dorsal pathway can build visual representations
894 of the spatial relations among distinct proto-objects almost
895 devoid of other visual attributes. In such representation with
896 an abstract quasi-conceptual content and almost no picto-
897 rial content, proto-objects are individuated as relata of spa-
898 tial relations. They become movable parts of visual scenes,
899 events or pictures and their respective spatial arrangement can
900 be subject to artistic composition by painters, designers or
901 architects.

902 One important feature of unilateral neglect is that neglect
903 patients are particularly vulnerable to the phenomenon of
904 extinction: if presented with two competing stimuli in their
905 contralesional left hemisphere, they will typically fail to per-
906 ceive the one further to their left. In other words, the stimulus
907 located more towards the ipsilesional side will extinguish its
908 competitor located more on the contralesional side. In one
909 experiment, [Driver and Vuilleumier \(2001\)](#) presented a ne-
910 glect patient with two conditions. In one condition, the stim-
911 ulus was a Kanizsa white square whose subjective contours
912 arose from the removal of the relevant quarter-segments from
913 four black circles. In the other condition, the stimulus con-
914 sisted of the four black circles in the same spatial positions,
915 but the formation of the subjective contours of the Kanizsa
916 white square was prevented by the fact that the four black cir-
917 cles were presented in their entirety. The patient extinguished
918 most left-sided presentations of the stimulus in bilateral tri-
919 als when the full presentation of the four black circles pre-
920 vented the formation of the subjective contours of the Kanizsa
921 white square. But extinction was much weaker when the pa-
922 tient could see the Kanizsa white square. In other words,
923 the neglect patient found it easier to allocate her perceptual
924 attention to one big object than to four competing smaller
925 objects.

926 The importance of this finding lies in the fact that in ne-
927 glect patients, the visual attributes of objects located in the
928 neglected hemispace are still covertly processed by the rel-
929 evant areas in the ventral pathway. But the patient remains
930 unaware of the visual attributes of stimuli located in their
931 neglected hemispace. By losing visual awareness of the rel-
932 ative spatial locations of objects in their neglected side, ne-
933 glect patients also lose visual awareness of their other visual
934 attributes of these objects. Loss of awareness of the spatial
935 relations between objects (provoked by a lesion in the right
936 inferior parietal lobe) produces loss of awareness of other
937 visual attributes. But the dependency seems asymmetrical:

938 loss of awareness of such visual properties of objects as their
939 colors, shapes, sizes or orientations does not seem to lead to
940 unawareness of the relative locations of objects.

941 On one hand, the claim that visual awareness of visual
942 attributes (such as color, shape, size and orientation) asym-
943 metrically depends on awareness of spatial relations among
944 objects is consistent with the view that the representation of
945 spatial relations among proto-objects has a quasi-conceptual
946 character. On the other hand, this asymmetrical dependency
947 fits with a conceptual analysis of what is the deep nature of
948 visual perception. Visual awareness of the size, shape and
949 orientation of one object consists in the perceptual compar-
950 ison between its relative size, shape and orientation and
951 those of neighboring objects. In other words, visual aware-
952 ness must satisfy the constraint of contrastive identification
953 (see Jacob & Jeannerod, 2003). But comparative perceptual
954 processing of the relative sizes, shapes and orientations of
955 two or more objects in turn presupposes the representation
956 of their relative spatial positions in some allocentric frame
957 of reference and the possibility to mentally manipulate this
958 representation.

959 Besides the case of neglect, further arguments in favor
960 of the asymmetrical dependency of visual awareness of ob-
961 ject identification upon awareness of spatial relations can be
962 drawn from the observation of patients presenting other types
963 of visuospatial disorders. Indeed, the processing of the spa-
964 tial orientation of an object may interfere with the visual
965 recognition and/or identification of that object. Warrington
966 and Taylor (1973) presented right parietal brain-lesioned pa-
967 tients with photographs of common objects (e.g., a basket)
968 taken from a non-conventional (or non-canonical) point of
969 view. The patients failed to recognize these objects, although
970 they had no problem recognizing the same objects when pre-
971 sented in a canonical view. Thus, the inability to mentally
972 manipulate spatial relationships of visual objects (e.g., by
973 mental rotation) might be responsible for the recognition im-
974 pairment. Another related condition is dorsal simultagnosia.
975 Typically, a dorsal simultagnosic patient will recognize most
976 objects but will be unable to see more than one at a time,
977 irrespective of their size. As a consequence of this condi-
978 tion, such patients cannot count objects; their description of
979 complex scenes is slow and fragmentary; they behave like
980 blind people when moving in a visual environment, groping
981 for things and bumping into obstacles. Dorsal simultagnosia
982 has been interpreted as a disorder of visual attention. Along
983 with Posner, Walker, Friedrich, and Rafal (1984) hypothesis,
984 Farah (1995) considers the possibility of a specific deficit in
985 *disengaging* one's visual attention: in order to be able to en-
986 gage one's visual attention onto a new stimulus, one must first
987 disengage one's visual attention from its prior and/or current
988 location. Parietal lobes would play a critical role in this at-
989 tentional mechanism. People with a bilateral parietal lesion
990 should thus present a 'sticky' attention on the current object
991 without the possibility to shift to another one and, by way of
992 consequence, without the possibility to build coherent spatial
relationships between them.

6. Conclusion 993

994 Visual cognition appears to be far more complex than pre-
995 viously suggested by the current models opposing either vi-
996 sual object perception and space perception or perception
997 and action. These models, which originated from the dou-
998 ble dissociation paradigm, attempted to match a given aspect
999 of visual function onto a given anatomical subdivision of
1000 the cortical visual system. The double dissociation paradigm,
1001 however, appears to be of a limited value when the number
1002 of the terms of the dissociation is greater than two. Clearly,
1003 according to the review above, there are more than two kinds
1004 of human visual representations and more than two-visual
1005 systems in the human brain.

1006 In this paper, we have argued that no valid comparison
1007 between visual representations can arise unless provision is
1008 made for three critical properties: their direction of fit, their
1009 direction of causation and the level of their cognitive or con-
1010 ceptual content. The cognitive (or conceptual) content in turn
1011 is a function of the level of processing. Representations aris-
1012 ing from earlier stages of processing—whether percepts or
1013 visuomotor representations—have very little or no concep-
1014 tual content. Elementary visual percepts, for example, arise
1015 from the automatic stage of semantic processing whereby
1016 basic visual attributes of an object are assembled and bound
1017 together. Low-level visuomotor representations of targets of
1018 prehension result from the automatic process of the visu-
1019 motor transformation. Whether they result from semantic or
1020 from pragmatic processing, higher order representations get
1021 their conceptual content from the connections between vi-
1022 sual cognition and other parts of the human cognitive system
1023 (such as the planning of action and semantic memory). They
1024 often proceed under conscious control. The two other criti-
1025 cal properties of visual representations are their mind/world
1026 direction of fit and their mind/world direction of causation.
1027 The output of the semantic processing of visual inputs has a
1028 full mind-to-world direction of fit and a full world-to-mind
1029 direction of causation: on one hand, it visually registers the
1030 way the world is, on the other hand, it is caused by what it
1031 represents. The output of the pragmatic processing of visual
1032 inputs has both a hybrid direction of fit and a hybrid direction
1033 of causation: it yields information for the benefit of intentions,
1034 which clearly have both a world-to-mind direction of fit and a
1035 mind-to-world direction of causation. An intention is both the
1036 representation of a goal and a cause of the transformation of
1037 a goal into a fact. This two-fold distinction segregates repre-
1038 sentations specialized for perception from those specialized
1039 for action.

1040 Since it penetrates deeply into visual knowledge of the
1041 world, visual perception cannot be limited to selecting an ob-
1042 ject from its surroundings, identifying it and giving it mean-
1043 ing. Semantic processing of visual inputs also implies compar-
1044 ison, which in turn requires that several objects be simul-
1045 taneously represented and analyzed: hence, object perception
1046 in turn presupposes the representation of spatial relationships
1047 among two or more objects in a coordinate system indepen-

dent from the perceiver. Spatial relationships in themselves carry cues for attributing meaning to an object, so that their processing is actually part of semantic processing of visual information. Thus, one has to consider that perception itself is actually distributed over the two classical pathways of the human visual system: the ventral and the dorsal cortical pathways. An intact (right) inferior parietal lobule is thus required for coding spatial relationships among objects in an allocentric frame of reference, which is itself part of the general process of semantic processing.

The above considerations about the role of the dorsal pathway in visual perception raise the point of the several functions of the human parietal lobe. Obviously, this brain region cannot be considered as a unitary entity with a single function. The superior parietal lobule carries visuomotor processing, a non-lateralized process common to monkey and man. The right inferior parietal lobule contributes to the perception of spatial relationships, a process with a mind-to-world direction of fit and a world-to-mind direction of causation. Finally, the left inferior parietal lobule contributes to still another type of representation, related to visually goal-directed action, i.e., with both a hybrid direction of fit and a hybrid direction of causation. The latter two processes are unique to man.

The identification of higher level motor representations as part of visual cognition—a critical aspect of pragmatic processing which has been one of the major themes of this paper—sheds new light on the relations between perception and action. As one moves from the automatic visuomotor transformation to the planning and execution of more complex actions involving the use and manipulation of tools, the distinction between action and perception loses much of its significance. On one hand, the ability to perform actions by using cultural tools and to pantomime the use of imaginary tools requires the storage of, and the ability to retrieve, schemas for tool manipulation. Such schemas are in turn formed on the basis of the observation of the actions of others. Thus, the ability to perform complex actions with tools relies on visual perception. On the other hand, schemas for the manipulation of tools are not only necessary for the appropriate actual or pantomimed use of tools but also for extracting the meaning of actions involving the use of tools by other agents. Not only is what one can do shaped by what one perceives, but also conversely what one can do shapes what one can perceive.

Not all perceived human actions are object-oriented actions. Nor are all human actions visually processed for the purpose of learning how to prepare and execute an action. Many human actions are directed towards conspecifics and many such actions are visually processed for the purpose of understanding their social and emotional significance. The processing of human actions directed towards conspecifics also gives rise to representations with various levels of conceptual content. At the lowest level of conceptual content is the visual perception of so-called “biological motion”. Then at various higher levels of visual processing are perceptual representations of emotions carried by facial movements and

expressions. They in turn provide visual information to the human mindreading system (in Baron-Cohen’s, 1995 terms), which underlies the attribution and recognition of mental states to others and to oneself.

Uncited references

Farah (1990) and Mountcastle, Lynch, Georgopoulos, Sakata, and Acuna (1975).

References

- Anscombe, G. E. (1957). *Intention*. Oxford: Blackwell.
- Baron-Cohen, S. (1995). *Mindblindness, an essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Berhmann, M., Winocur, G., & Moscovitch, M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature*, *359*, 636–637.
- Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex*, *14*, 129–133.
- Bridgeman, B., Kirsch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics*, *29*, 336–342.
- Clark, M. A., Merians, A. S., Kothari, A., Poizner, H., Macauley, B., Rothi, L. J. G., et al. (1994). Spatial planning deficits in limb apraxia. *Brain*, *117*, 1093–1106.
- Decety, J., & Jeannerod, M. (1996). Fitts’ law in mentally simulated movements. *Behavioral Brain Research*, *72*, 127–134.
- Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., et al. (1994). Mapping motor representations with PET. *Nature*, *371*, 600–602.
- Dretske, F. (1969). *Seeing and knowing*. Chicago: Chicago University Press.
- Dretske, F. (1978). Simple seeing. In F. Dretske (Ed.), *Perception, knowledge and belief*. Cambridge: Cambridge University Press.
- Driver, J., & Vuilleumier. (2001). Perceptual awareness and its loss to unilateral neglect and extinction. In S. Dehaene & L. Naccache (Eds.), *The cognitive neuroscience of consciousness*. Cambridge, MA: MIT Press.
- Faillenot, I., Decety, J., & Jeannerod, M. (1999). Human brain activity related to the perception of spatial features of objects. *Neuroimage*, *10*, 114–124.
- Farah, M. J. (1990). *Visual agnosia. Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.
- Farah, M. J. (1995). Current issues in the neuropsychology of image generation. *Neuropsychologia*, *33*, 1455–1471.
- Farah, M. J., Hammond, K. M., Levine, D. M., & Calvanio, R. (1988). Visual and spatial mental imagery. Dissociable systems of representation. *Cognitive Psychology*, *20*, 439–462.
- Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: A PET study. *Neuroimage*, *18*, 324–333.
- Géradin, E., Sirigu, A., Lehericy, S., Poline, J. B., Gaymard, B., Marsault, C., et al. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex*, *10*, 1093–1104.
- Glover, S. Separate visual representations in the planning and control of action. *Behavioral and Brain Sciences*, in press.
- Goldenberg, G., Hartmann, K., & Schlott, I. (2003). Defective pantomime of object use in left brain damage: Apraxia or asymbolia? *Neuropsychologia*, *41*, 1565–1573.
- Goldenberg, G., Müllbacher, W., & Nowak, A. (1995). Imagery without perception. A case study of anosognosia for cortical blindness. *Neuropsychologia*, *33*, 1373–1382.

- 1163 Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991).
1164 A neurological dissociation between perceiving objects and grasping
1165 them. *Nature*, *349*, 154–156.
- 1166 Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Lo-
1167 calization of grasp representations in humans by PET: 2. Observa-
1168 tion compared with imagination. *Experimental Brain Research*, *112*,
1169 103–111.
- 1170 Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini,
1171 P., & Grady, C. L. (1994). The functional organization of human
1172 extrastriate cortex. A PET-rCBF study of selective attention to faces
1173 and locations. *Journal of Neurosciences*, *14*, 6336–6353.
- 1174 Ingle, D. J. (1973). Two visual systems in the frog. *Science*, *181*,
1175 1053–1055.
- 1176 Jacob, P., & Jeannerod, M. (2003). *Ways of seeing. The scope and limits*
1177 *of visual cognition*. Oxford: Oxford University Press.
- 1178 Jeannerod, M. (1986). The formation of finger grip during prehension. A
1179 cortically mediated visuomotor pattern. *Behavioural Brain Research*,
1180 *19*, 99–116.
- 1181 Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford:
1182 Blackwell.
- 1183 Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism
1184 for motor cognition. *Neuroimage*, *14*, S103–S109.
- 1185 Jeannerod, M., & Frak, V. G. (1999). Mental simulation of action in
1186 human subjects. *Current Opinions in Neurobiology*, *9*, 735–739.
- 1187 Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of grasping
1188 movements following a bilateral posterior parietal lesion. *Neuropsychy-*
1189 *chologia*, *32*, 369–380.
- 1190 Johnson, S. H. (2000). Thinking ahead: The case of motor imagery in
1191 prospective judgements of prehension. *Cognition*, *74*, 33–70.
- 1192 Johnson, S. H., Rotte, M., Grafton, S. T., Hinrichs, H., Gazzaniga, M. S.,
1193 & Heinze, H. J. (2002). Selective activation of a parietofrontal circuit
1194 during implicitly imagined prehension. *Neuroimage*, *17*, 1693–1704.
- 1195 Köhler, S., Kapur, K., Moscovitch, M., Winocur, G., & Houle, S. (1995).
1196 Dissociation of pathways for object and spatial vision. A PET study
1197 in humans. *NeuroReport*, *6*, 1865–1868.
- 1198 Kosslyn, S. M., Alpert, N. M., Thomson, W. L., Maljkovic, V., Wiese,
1199 S. B., Chabris, C. F., et al. (1993). Visual mental imagery activates
1200 topographically organized visual cortex: PET investigation. *Journal of*
1201 *Cognitive Neuroscience*, *5*, 263–287.
- 1202 Kosslyn, S. M., DiGirolamo, G. J., Thompson, W. L., & Alpert, N.
1203 M. (1998). Mental rotation of objects versus hands: Neural mech-
1204 anisms revealed by positron emission tomography. *Psychophysiology*,
1205 *35*, 151–161.
- 1206 Levine, D. N., Warach, J., & Farah, M. (1985). Two visual systems
1207 in mental imagery: Dissociation of “what” and “where” in imagery
1208 disorders due to bilateral posterior cerebral lesions. *Neurology*, *35*,
1209 1010–1018.
- 1210 Marshall, J., & Halligan, P. W. (1994). The yin and yang of visuospatial
1211 neglect. A case study. *Neuropsychologia*, *32*, 1037–1057.
- 1212 Millikan, R. G. (1996). Pushmi-pullyu representations. In J. Tomberlin
1213 (Ed.), *Philosophical perspectives: vol. IX*. Atascadero, CA: Ridgeview
1214 Publishing.
- 1215 Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*.
1216 Oxford: Oxford University Press.
- 1217 Milner, A. D., Paulignan, Y., Dijkerman, H. C., Michel, F., & Jean-
1218 nerod, M. (1999). A paradoxical improvement of misreaching in op-
1219 tic ataxia: New evidence for two separate neural systems for visual
1220 localization. *Proceedings of the Royal Society of London. Series B*,
1221 *266*, 2225–2229.
- 1222 Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H., & Acuna,
C. (1975). Posterior parietal association cortex of the monkey: Com-
mand functions for operations within extra-personal space. *Journal of*
Neurophysiology, *38*, 871–908.
- 1223 Nilsson, L. -G., Nyberg, L., Klingberg, T., Aberg, C., Persson, J., &
1224 Roland, P. E. (2000). Activity in motor areas while remembering
1225 action events. *NeuroReport*, *11*, 2199–2201.
- 1226 Ochipa, C., Rapsack, S. Z., Maher, L. M., Rothi, L. J. G., Bowers,
1227 D., & Heilman, K. M. (1997). Selective deficit of praxic imagery in
1228 ideomotor apraxia. *Neurology*, *49*, 474–480.
- 1229 Parsons, L. M. (1994). Temporal and kinematic properties of motor
1230 behavior reflected in mentally simulated action. *Journal of Exper-*
1231 *imental Psychology. Human Perception and Performance*, *20*, 709–
1232 730.
- 1233 Parsons, L. M., Fox, P. T., Downs, J. H., Glass, T., Hirsch, T. B., Martin,
1234 C. C., et al. (1995). Use of implicit motor imagery for visual shape
1235 discrimination as revealed by PET. *Nature*, *375*, 54–58.
- 1236 Pinker, S. (1985). Visual cognition: An introduction. In S. Pinker (Ed.),
1237 *Visual cognition* (pp. 1–65). Cambridge, MA: MIT Press.
- 1238 Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984).
1239 Effects of parietal lobe injury on covert orienting of visual attention.
1240 *Journal of Neuroscience*, *4*, 1863–1874.
- 1241 Pylyshyn, Z. (2000a). Situating vision in the world. *Trends in Cognitive*
1242 *Sciences*, *4*(5), 197–207.
- 1243 Pylyshyn, Z. (2000b). Visual indexes, preconceptual objects and situated
1244 vision. *Cognition*, *80*, 127–158.
- 1245 Rosenbaum, D. A., & Jorgensen, M. J. (1992). Planning macroscopic
1246 aspects of motor control. *Human Movement Science*, *11*, 61–69.
- 1247 Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking
1248 during simulation of action: A PET investigation of agency. *Nature*
1249 *Neurosciences*, *4*, 546–550.
- 1250 Schneider, G. E. (1969). Two visual systems. *Science*, *163*, 895–902.
- 1251 Searle, J. R. (1983). *Intentionality, an essay in the philosophy of mind*.
1252 Cambridge: Cambridge University Press.
- 1253 Servos, P., & Goodale, M. A. (1995). Preserved visual imagery in visual
1254 form agnosia. *Neuropsychologia*, *33*, 1383–1394.
- 1255 Shepard, R. N., & Metzler, J. (1971). Mental rotation of three dimensional
1256 objects. *Science*, *171*, 701–703.
- 1257 Shubotz, R. I., & von Cramon, Y. (2002). Predicting perceptual events
1258 activates corresponding motor schemes in lateral premotor cortex: An
1259 fMRI study. *Neuroimage*, *15*, 787–796.
- 1260 Sirigu, A., & Duhamel, J. R. (2001). Motor and visual imagery as two
1261 complementary but neurally dissociable mental processes. *Journal of*
1262 *Cognitive Neuroscience*, *13*, 910–919.
- 1263 Sirigu, A., Duhamel, J. R., Cohen, L., Pillon, B., Dubois, B., & Agid, Y.
1264 (1996). The mental representation of hand movements after parietal
1265 cortex damage. *Science*, *273*, 1564–1568.
- 1266 Sirigu, A., Cohen, L., Duhamel, J. R., Pillon, B., Dubois, B., & Agid, Y.
1267 (1995). A selective impairment of hand posture for object utilization
1268 in apraxia. *Cortex*, *31*, 41–55.
- 1269 Stelmach, G. E., Castiello, U., & Jeannerod, M. (1994). Orineting the fin-
1270 ger opposition space during prehension movements. *Journal of Motor*
1271 *Behavior*, *26*, 178–186.
- 1272 Ungerleider, L., & Mishkin, M. (1982). Two cortical visual systems. In
1273 D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of*
1274 *visual behavior* (pp. 549–586). Cambridge: MIT Press.
- 1275 Ungerleider, L. G., & Haxby, J. V. (1994). ‘What’ and ‘where’ in the
1276 human brain. *Current Opinion in Neurobiology*, *4*, 157–165.
- 1277 Tomasino, B., Rumiati, R. I., & Umiltà, C. A. (2002). Selective deficit of
1278 motor imagery as tapped by a left–right decision of visually presented
1279 hands. *Brain and Cognition*.
- 1280 Warrington, E. K., & Taylor, A. M. (1973). The contribution of right
1281 parietal lobe to object recognition. *Cortex*, *9*, 152–164.
- 1282
1283