UNIVERSITY OF URBINO
Carlo Bo
Department of Pure and Applied Science
Doctoral Programme in Earth Sciences
and Complexity Science –
Curriculum in Complexity Science
XXVIII Cycle

Scientific Field: M/FIL02 Logic and
Philosophy of Science

UNIVERSITY OF ANTWERP
Department of Philosophy
Doctoral Programme in Philosophy
(Centre for Philosophical Psychology,
Between Perception and Action)

Perception, Action and Neuroscience

SUPERVISORS: Professor **Mario Alai**
Professor **Bence Nanay**

Ph.D. STUDENT: Gabriele Ferretti

ACADEMIC YEAR
2014/2015

To Genovino, who lent me the car.
To Manuela, who accompanied me,
with incommensurable love
and devotion.
To Gabriele, who provided the fuel.
To Mario, who assisted me patiently, carefully
and unhesitatingly
during my research for action possibilities.
To Bence, who believed in me
and showed me new, marvelous,
motor representations.

Abstract in English - Perception, Action and Neuroscience

The main aim of this thesis is to offer new insights in our understanding of the relationships between visual perception and action. Starting from experimental results in vision and motor neuroscience, it will do this by suggesting new functional characteristics – which are neglected in the literature – of vision-for-action and of motor perception. The crucial task of these characteristics is to compute, from a motor point of view, the most suitable way we can interact with the external environment. A better understanding of these characteristics requires an investigation into the nature of the computational mechanisms through which our visuomotor brain can lead us to perceive the possibilities of action in this external environment.

Describing in a coherent way the nature of the relation between the visual processes that allow us to guide action and the motor processes at the basis of our motor skills is one of the most yarned enterprises in both contemporary philosophy and neuroscience, insofar as opening the "black-box" of the mental representations by which vision-for-action is subserved – in the literature they are called motor representations (MRs) – means to understand in general how the way embodied agents can actively couple with the external environment.

Here, I want to offer a specific contribution to several open, central issues, each concerning the relation between visual perception and action, and each familiar from the literature concerning the philosophy of neuroscience.

Starting from two specific empirical frameworks, the one of the 'Two Visual Systems Model' and the other of a model of 'Motor Perception', the project will shed new light on the mental processes that allow us to visually perceive action possibilities in the external environment. Indeed, it is known that these processes convert the visual information about the objects we are faced with into motor information that we can use to interact with those objects. But what kind of mental representations do these processes rely on? This is the foundational question guiding my overall research. Answering this general question requires a discussion of several more specific issues connected with it. I will confront these issues, and thus answer my main question, by starting from the philosophical analysis of the processing of the cortical anatomo-functional correlates, as well as the computational mechanisms these mental phenomena – that is, MRs – rely on.

Here is a brief overview of the chapters of this thesis, and the related issues I focus on. Chapter 1 introduces the topic. Chapter 2 reviews and synthesises the massive sets of neuroscientific evidence in a coherent philosophical theory of MRs, something we lack at the moment. The goal of chapter 3 is to show that emotions play a crucial role in forming our mental antecedents of action, insofar as they assist vision-for-action from its early stage processing. Chapter 4 aims to defend the idea, completely new, that the representation of action properties, which is a function of MRs, is possible not only for normal objects, but also for depicted objects. Chapter 5 suggests that depicted objects cannot foster the visual feeling of presence because it is strictly linked to the perception of absolute depth cues that give rise to qualitatively rich stereopsis, which is linked to the perception of the possibility of reliable motor interaction with the object we face with.

Abstract in Dutch - Waarneming, Handeling, en Neurowetenschap

Het hoofddoel van dit proefschrift is nieuw inzicht te verschaffen in ons begrip van de verhouding tussen visuele waarneming en handeling. Het doet dit door nieuwe functionele eigenschappen van visie-voor-handelen en motorwaarneming aan te dragen, uitgaande van experimentele bevindingen in de visuele en neurowetenschappen. De hoofdtaak van deze eigenschappen is het bepalen, bezien vanuit een motor-perspectief, van de meest geschikte vorm van interactie met de externe omgeving. Deze eigenschappen zijn in de huidige discussie nog niet adequaat behandeld. Een beter begrip van deze eigenschappen vereist een onderzoek naar de aard van de computationele mechanismen waarmee onze visuomotorhersenen ons in staat stellen handelingsmogelijkheden in de externe omgeving waar te nemen.

Een coherente beschrijving van de aard van het verband tussen enerzijds de visuele processen die ons in staat stellen handeling te sturen, en de motorprocessen aan de basis van onze motorische vaardigheden anderzijds, is vandaag één van de grootste struikelblokken zowel voor de wijsbegeerte als voor de neurowetenschappen. Dit komt omdat het openen van de 'zwarte doos' van mentale representaties die visie-voor-handelen dienen—de literatuur spreekt van 'motor representaties' (MRs)—een algemeen begrip vergt van de manier waarop belichaamde actoren zich actief kunnen koppelen met een externe omgeving.

Ik zal een bijdrage leveren aan enkele openstaande centrale vraagstukken, bekend uit literatuur over de filosofie van de neurowetenschappen, over de verhouding tussen visuele waarneming en handelen.

Het project zal uitgaan van twee specifieke empirische raamwerken, het ene het 'Two Visual Systems Model' en het andere een 'Motor Perception'-model, en zal op basis hiervan nieuw licht te werpen op de mentale processen die ons in staat stellen om handelingsmogelijkheden in de externe omgeving te zien. Het is inderdaad bekend dat deze processen de visuele informatie over de objecten waar we tegenover staan omzetten naar motorinformatie die we kunnen aanwenden in interactie met deze objecten. Maar op wat voor soort mentale representaties berusten deze processen? Dit is de fundamentele vraag die mijn onderzoek stuurt. Het beantwoorden van deze meer algemene vraagstelling vereist discussie van een aantal specifieke problemen die ermee verbonden zijn. Ik zal deze specifieke problemen behandelen, en zodoende mijn hoofdvraag beantwoorden, door te beginnen met een filosofische analyse van de verwerking in de corticale anatomo-functionele correlaten, alsmede de computationele mechanismen waarop deze mentale fenomenen – MRs – berusten.

Hier is een kort overzicht van de hoofdstukken van dit proefschrift, en de relevante problemen waar ik me op richt.

Hoofdstuk 1 leidt het onderwerp in. Hoofdstuk 2 brengt de zeer omvangrijke verzameling van neurowetenschappelijke data samen in een coherente filosofische theorie van motor-representaties, iets dat tot op heden niet voorhanden was. Het doel van Hoofdstuk 3 is te laten zien dat emoties een cruciale rol spelen in het vormen van de mentale voorboden van handeling, in zoverre ze visie-voor-handelen assisteren vanaf de vroegste stadia van verwerking. Hoofdstuk 4 richt zich op de verdediging van de volledig

nieuwe stelling dat de representatie van handelingseigenschappen, welk een functie is van motor-representaties, niet alleen mogelijk is voor normale objecten, maar tevens voor afgebeelde objecten. Hoofdstuk 5 stelt dat afgebeelde objecten niet het gevoel van aanwezigheid aan kunnen moedigen, aangezien dit gevoel strikt gekoppeld is aan de waarneming van absolute diepte-*cues* die leiden tot kwalitatief rijke stereopsis, welke gekoppeld is aan de waarneming van de mogelijkheid tot betrouwbare motor-interactie met het object waar we tegenover staan.

Contents

Acknowledgements

This work is the result of three years of study concerning the relationships between perception and action. Most of the ideas exposed in this thesis have been presented at the following international conferences: in 2015, at the international conference on Model-Based Reasoning in Science and Technology in Sestri Levante, University of Pavia, Italy; in 2014, at the AISC International Conference for Cognitive Sciences, University of Rome 3, at the European Society for Philosophy and Psychology at the University of Messina in Noto, at the Annual Conference for Consciousness and Experiential Psychology in Cambridge, at the International Conference for Analytic Philosophy at the University of L'Aquila, at the European Conference for Analytic Philosophy at the University of Bucharest, at the International Conference for Logic and Philosophy of Science at the University of Rome 3; in 2013 at the Salzburg Conference for Analytic Philosophy.

I am grateful to those people who actively participated in the discussions in these occasions. A lot of those ideas have also been presented in four different departmental colloquia, three at the department of Pure and Applied Science of Urbino, during those workshops organized by the research group of "Complexity Science" and one at the Universiteit Antwerpen, during a workshop organized with the research group "Between Perception and Action".

A special thank goes to Bence Nanay, Mario Alai, Gino Tarozzi, Adriano Angelucci, Silvano Zipoli Caiani, Anna Maria Borghi, Pierre Jacob, Corrado Sinigaglia, Alfredo Paternoster, Michele Di Francesco, Pierluigi Graziani, Vincenzo Fano, Claudio Calosi, Andrea Borghini, Chiara Brozzo, Dan-Cavendon Taylor, Joseph Brenner, Mohan Matthen, Achille Varzi, Maarten Steenhagen, Laura Gow, Grace Helton, Kris Goffin, Nick Young, Neil Van Leeuwen, Peter Fazekas, Angelica Kauffman, Sara Matera, Nicola Bruno, Anna Maria Borghi, Luca Barlassina, Angel Del Pobil, Eris Chinellato, Michael Segre, Luciano Boi, Joshua Shepherd, Genovino Ferri, Mike Webb.

1 Introduction: The Jungle between Vision and Action

> The hand has a very complex anatomical structure. Functionally, movements of the hand require a coordinated interplay of the 39 intrinsic and extrinsic muscles acting on 18 joints. Among all the joints of the hand, of particular importance is the carpometacarpal joint of the thumb. This joint is of a saddle type and its immense significance for the hand function emanates from the extra mobility this joint is endowed with, resulting in the opposition of the thumb to the other fingers. The plethora of bones, joints, and muscles of which the hand is constituted gives to this structure amazing biomechanical complexity. From the kinematic perspective, the hand has over 20 degrees of freedom. Thus the question arises: how does the brain control the hand? (Raos et al. 2006: 709)

This work is about the relationships between vision and action. Starting from experimental results in vision and motor neuroscience, the main aim of this thesis is to offer new insights in our understanding of the relationships between visual perception and action, by suggesting new functional characteristics – which are neglected in the literature - of vision-for-action and of motor perception, the crucial task of which is to compute the most suitable way we can interact, from a motor point of view, with the external environment. That means to investigate the nature of the computational mechanisms through which our visuomotor brain can lead us to perceive the possibilities of action in this external environment.

The relationship between visual perception and action is old as history of knowledge is. All the greatest scholars involved in the study of the mind (or, in particular historical periods, of the soul) and of its access and relation to the external world have tried to understand the way we can perceive the external environment and, then, how we can act, on the basis of how our perception works, on what we face with. From the ancient greeks, Plato (see Remes 2014), Aristotle (see Corcilius 2014; Tuominen 2014) to the medieval scholars, such as Augustine (see Silva 2014a), Avicenna (see Kaukua 2014), Averroes (see Brenet 2014; Silva 2014b), passing through the scholars of modern age, such as Nicholas of Cusa, Hobbes (for a review see Leijenhorst 2014), Descartes (see Wee 2014), Locke (see Lähteenmäki 2014), Spinoza (see Viljanen 2014), Berkeley (see Paukkonen 2014), Descartes and Kant (Hatfield 2014) – for a complete historical review in which all these works are recollected see (Silva and Yrjönsuuri 2014; see also Lagerlund 2007) – until the phenomenological tradition, which has been crucial in explaining the role of the body for the relation between perception and action (Husserl 1907/1998, Merleau-Ponty 1962), one of the most important goal has been to explain the relation between our senses, our causal power on the external environment and the way our body shapes our relationship with the external world.

While the possible relationship between visual perception and action have been in the spotlight of the contemporary philosophical reflection (Gibson 1966, 1979; MacKay 1962, 1967, 1973; Merleau-Ponty 1968), the investigation of the specific nature of this relationship has been meticulously deepened, recently, thanks to the impressive amount of evidence straddling the divide between motor and vision neuroscience, with a particular philosophical interpretation (Borghi and Riggio 2015; Chinellato and Del Pobil 2015; Turella and Lignau 2014; Castiello 2005; Castiello and Begliomini 2008; Fadiga et al. 2000; Kandel et al. 2013, Chalupa and Werner 2004), as well as thanks to the philosophical reflection about these experimental results (Jeannerod 2006; Jacob and Jeannerod 2003; Rizzolatti and Sinigaglia 2008; Nanay 2013b; Grush 2004; Briscoe 2009; Briscoe and Grush 2014; Jacob 2005; Seth 2014, 2015; Friston et al. 2010, 2013; Clark 1999, 2001, 2002, 2009, 2007; Prinz et al. 2013).

Here, I want to give a specific contribution about specific open issues, concerning this relation between visual perception and action, with respect to the literature concerning the philosophy of neuroscience.

There are two important background specifications I have to address before starting with the overview of my work here. First, I want to specify that I will not take part to the contemporary divide in the literature between those who believe that perception is the construction of internal representations (Nanay 2010a, b, 2011a, b, 2012, 2013a, b, 2014a, b; Siegel 2006, 2010a, b, 2014; Campbell 2008; Martin 2008; Kelly 2008; Noë 2008a, b; Prosser 2011) and those who, contrarily, deny this idea (Hutto and Myin 2013; Hutto 2005; Chemero 2009; Noë 2004; O'Regan and Noë 2001; Hurley 1998, Varela Thompson and Rosch 1991; Findlay and Gilchrist 2003; Bishop and Martin 2014). It is well know, indeed, that, though considering perceptual states as representations has been considered the most natural way of describing our perceptual system, this framework has recently been questioned. Notwithstanding this, the balance leans on representationalism, to the extent that describing perceptual states as representations has some important explanatory advantages (Pautz 2010, Nanay 2013a, b, 2014b; Ferretti forthcoming, Ferretti and Alai forthcoming). I leave aside this debate. Indeed, as we shall see in each section, this work is written in terms of (perceptual) representations. That is, I endorse the general picture about the mind which also suggests a representational view of perception.

Second, some authors inquired the extent to which action is important for visual perception, assuming that vision is a form of action and that a theory of vision can be grounded on action (Noë 2004; O'Regan and Noë 2001). However, several critics have shown the inefficacy of this attempt (Jacob 2006, 2008; Jacob and Jeannerod 2003; Rowlands 2006, 2007; Albertazzi, van Tonder and Vishwanath 2010; Mandik 2005; Gangopadhyay and Kiverstein 2009; Gangopadhyay 2010; Pearson and Westbrook 2015). Since, saying that vision is a form of action could lead to different misunderstandings, I detach my work here from this idea; however, most of the empirical evidence I report in this work might be interpreted as going in this direction. Here, I am just investigating the different computational aspects of those computational resources of vision that are crucial in order to guide our action, without committing myself to a constitutive or a causal relation between perception and action. Also, I am not assuming, *pace* Gibson, that action guidance is the only scope vision has to carry out (for a review see Jacob and Jeannerod 2003: Ch. 6; Nanay 2013b).

Finally, there are different general ways of coupling philosophy with neuroscience (Bennett and Hacker 2003, 2012; Bennett et al. 2007; Bickle et. al 2013; Brooks and Akins 2005). On the one hand we can (a) use the tools of analytic philosophy in order to answer to genuine philosophical problems concerning the foundation of neuroscience, which neuroscience constitutively fails to take into account. On the other hand, we can (b) use analytic philosophy in order to analyze experimental evidence whose interpretation is not clear. Also, we might (c) use empirical evidence from neuroscience to have a good empirical background in order to strengthen a philosophical thesis that seems to be already properly defended with theoretical arguments – and, thus, the experimental result would be a further empirical confirmation of what we know as being the case following the philosophical arguments. However, sometimes (d) empirical evidence from neuroscience seems to play a crucial role in showing us which direction we have to pursue in order to sort out such and such philosophical issues the literature is divided about, leading us to some theoretical crossroad – at the same time, sometimes it is the neuroscientific reflection to be philosophy/theory-laden. In these occasions, empirical evidence - with a solid philosophical completion, though - can give us the answer to certain philosophical problems remained unsolved. Crucially, there are cases in which (e) important issues arise within the literature in philosophy of neuroscience, which are the result of the mix of the reflection from the two disciplines and cannot be accounted neither just from philosophy, nor just from neuroscience: they need the interplay of these two ways of looking at the problem.

The methodology of this work is primarily of the kind (e), though it also follows (c) and (d). So, before going in *medias res*, I have briefly to say something about the empirical framework my philosophical reflection starts from.

1.2 The Empirical Background I start from: Vision and Motor Neuroscience

Several contemporary debates involving both philosophy and neuroscience attempt to define the nature of the relation between visual perception and (vision-for-)action, with a special focus on space perception.

A very interesting idea is that, rather than propositional attitudes like beliefs and desires, what mediates between the sensory input of visual perception and the motor output of action is the presence of motor representations (MRs) involved in representing the action properties necessary for action performance (Nanay 2013b, 2014a; Jacob and Jeannerod 2003; Jeannerod 2006; Pacherie 2000; Butterfill and Sinigaglia 2012).

Describing in a coherent way the nature of the relation between the visual processes that allow us to guide action and the motor processes at the basis of our motor skills is one of the most yarned enterprises in both contemporary philosophy and neuroscience (Briscoe 2009), insofar as opening the "black-box" of the mental representations vision-for-action is subserved by, which in the literature are indeed called MRs, means to understand the way embodied agents can actively couple with the external environment.

However, the mental antecedents of action, that is MRs, are not always consciously accessible; thus, to investigate and naturalize them we must turn to neuroscience (Nanay 2014a). There are two major empirical frameworks we have to focus on if we want to investigate the relationships between visual perception and action by investigating MRs. The first one is the "Two Visual Systems Model" (TVSM) and the second one is the one about Motor Perception, which I illustrate in what follows.

1.2.1 The Two Visual Systems Model

A common ground to link visual perception with action is the TVSM, which suggests the presence, in humans and other mammals, of a separation of the visual pathways, grounded on distinct anatomo-functional structures (Milner and Goodale 1995): one for visual recognition, the ventral stream, and one for visually guided action, the dorsal stream. They can be dissociated due to cortical lesions. Lesions in the dorsal stream (the occipito-parietal network from the primary visual cortex to the posterior parietal cortex) impair one's ability to use what one sees to guide action (optic ataxia), but not object recognition; lesions in the ventral stream (the occipito-temporal network from the primary visual cortex to the infertemporal cortex) impair one's ability to recognize things in the visual world (visual agnosia), but not action guidance (Jacob and Jeannerod 2003). Moreover, we have behavioral studies of normal subjects involving visual illusions that can deceive the ventral stream but not the dorsal one; thus, it seems that, contrarily to ventral perception, dorsal perception is completely impenetrable by consciousness. As we shall see, this is not the whole story about the relationships between the ventral and the stream and, ipso facto, about the relationships between conscious perception and vision-for-action (Schenk and MacIntosh 2010; Bruno and Battaglini 2008; McIntosh and Schenk 2009; Briscoe 2009; Gangopadhyay, Madary and Spencer 2010; Jacob 2008, 2014, 2015; Jacob and de Vignemont 2010; Jeannerod and Jacob 2005).

1.2.2 (Visuo-)Motor Perception

For decades the functions of motor areas have been reduced to executive tasks. This framework found the sensory areas in occipital lobe and motor areas in posterior part of frontal lobe. Between them, associative (temporo-parietal) areas put together information from sensory areas and send percepts to motor areas to organize movement: the idea was that associative areas were committed to higher cognitive functions, while motor areas dealt only whit motor execution. But this framework was not able to explain the way the motor system can translate visual perception in movement. Surprisingly, contemporary evidence suggested that while the motor system is not only involved in executive functions, the motor cortex influences the perceptual side of the parietal lobe which is strongly involved in motor activities so to be considered a part of the motor brain (Mountcastle et al. 1975; Fogassi and Luppino 2005; Rizzolatti and Sinigaglia 2008).

This idea has been confirmed by several experimental results about the existence of visuomotor phenomena, in the dorsal stream, responsible for the translation of object features in motor commands, involved in the visual perception of action possibility in peripersonal spatial coordinates, in which a distinction between the visual and the motor aspect of the process is constitutively indistinguishable, insofar the informational switch from the visual input to the motor output is functionally fuzzy and hardly detectable; a clear example of motor perception (Fadiga et al. 2000; Fogassi and Luppino 2005; Rizzolatti and Sinigaglia 2008; Castiello 2005, Turella and Lignau 2014, Romero et al. 2014, Raos et al. 2006, Gallese 2007, Rizzolatti and Matelli 2003; Graziano 2009; Rizzolatti 2000; Gazzaniga 2000; Jacob 2005; Borghi and Riggio 2015; Jacob and Jeannerod 2003, Jeannerod 2006; Romero et al. 2014). Furthermore, there are several experimental results - and several philosophical theories based on these results - suggesting the crucial importance of the motor system for our perception of action (Sinigaglia 2013; Sinigaglia and Butterfill 2015a, b; Gallese and Sinigaglia 2011a, b, c, 2013, 2014; Rizzolatti and Sinigaglia 2008, 2013).

These are the empirical frameworks my philosophical analysis starts from. I'll use some important results coming from these two empirical frameworks, and use some of the different methodologies I reported in (§1), in order to propose my arguments. Now, I offer an overview of the philosophical issues I'm going to focus on.

1.3 Visual Perception in Action: An Analytic Overview

Starting from these two empirical frameworks, the main aim of this project is to shed new light on the mental processes that allow us to visually perceive action possibilities in the external environment by converting the visual information about the objects we face with, in motor information we can use to interact with them. What kind of mental representations do these processes rely on? This is the foundational question guiding my overall research. However, several particular issues are related to this more general question. I try to confront with these issues, and thus to answer this question, by starting from the philosophical analysis of the processing of the cortical anatomo-functional correlates, as well as the computational mechanisms these mental phenomena - that is MRs - rely on.

Here's a brief analytical overview of the chapters of this thesis, and the related issues I focus on.

1.3.1 Through the Forest of Motor Representations

I said that the mental antecedents of action are not always consciously accessible; thus, to investigate and naturalize them we must turn to neuroscience (Nanay 2014a). Following neuroscience, several philosophers assume that there is a single representational mechanism lying in (and mediating) between the visual and motor processes involved in different functions for shaping suitable action performances: a MR (see Nanay 2013b; Butterfill and Sinigaglia 2012). The problem in the literature that motivates my project is the lack of common agreement about both the functions and the neural underpinnings of MRs. Hence, in the second chapter of this work, I suggest a solution by developing a non-monolithic view of MRs and arguing that a MR is the ensemble of different sub-representational phenomena, each one with a different function and a precise neural underpinning in relation to the TVSM. Accordingly, I recollect the massive sets of neuroscientific evidence in a coherent philosophical theory of MRs, something we lack at the moment.

1.3.2 Motor Representations and Emotions

Another goal I want to pursue is to investigate the role of emotions in action performance, which is not considered by most authoritative action theorists (see Pacherie 2002), as well as by the literature on MRs. In chapter 3, I suggest that the representation of the emotionally relevant properties of objects (being pleasant, dangerous, etc.) can influence, inhibiting or eliciting, the way we represent the action properties/possibilities (being graspable, climbable, etc.) of each object. I strengthen this argument with neural evidence that the neural underpinnings of MRs are deeply interconnected with emotional areas (Barrett and Bar 2009) and behavioral evidence that while neutral graspable objects approached without any risk activate a facilitating motor response, dangerous objects activate aversive motor responses, generating an interference-effect (Anelli, Borghi and Nicoletti 2012). My goal is to show that emotions play a crucial role in forming our mental antecedents of action, insofar as they assist vision-for-action from its early stage processing.

1.3.3 Pictures, Action Properties and Motor Related Effects

In chapter 4 I defend the idea, completely new, that the representation of action properties, which is a function of MRs, is possible not only with normal objects, but also with depicted objects. This idea is, following the TVSM (Nanay 2011a, 2015; Matthen 2010), commonly denied in the philosophical literature, because depicted objects are not physically manipulable and egocentrically localized, as normal objects, by dorsal perception. I want to suggest that, though depicted objects cannot be egocentrically localized as normal objects, face-to-face perception and the perception of depicted objects share the same subpersonal dorsal representation of action properties, and several sets of evidence are on my side (Romero et al. 2014; Chao and Martin 2000). This is possible only when the picture vehicle/surface falls within the peripersonal space of the observer (Borghi and Riggio 2015), because dorsal perception cannot distinguish between depicted and normal objects (Westwood et al. 2002). So, I defend the idea that, once an object, whether depicted or real, is perceived as apparently located in the peripersonal space of the observer, dorsal perception responds. Of course, this does not mean that we can act upon depicted objects, but just that our subpersonal dorsal representations are active when we visually perceive them. A further crucial implication of my new proposal is that, if also dorsal perception responds to depicted objects, then, picture perception and face-to-face perception are very similar perceptual states.

Issues in (§1.3.1, chapter 2), (§1.3.2, chapter 3) and (§1.3.3, chapter 4) are not distinct enterprises, but different parts of a comprehensive theory of MRs.

1.3.4 Visual Feeling of Presence

As argued in (§1.3.3, concerning the chapter 4) - both normal and depicted objects can be represented as offering action properties; thus, face-to-face and picture perception are not so different as concerns the visual system processing. Nevertheless, is indisputable that normal objects are perceived as present, at the visual level, in a way depicted objects cannot (Matthen 2010; Nanay 2015). According to the literature, depicted objects cannot foster the visual feeling of presence (FOP) because they cannot be represented in egocentric coordinates. In chapter 5 I strengthen this idea, but aims to specify that egocentric localization is responsible for visual FOP because it is strictly linked to the perception of absolute depth cues that give rise to the most important aspect of visual FOP: qualitatively rich stereopsis (Vishwanath 2014) which is linked to the perception of the possibility of reliable motor interaction with the object we face with. In relation to (§1.3.3, chapter 4), here I suggest that only normal objects can foster in us the conscious perception of reliable motor interaction, for which the subpersonal dorsal representations of action properties are not enough. I also aim to suggest that, contrarily to the received view about feelings (Dokic 2012), visual FOP is not the result of a metacognitive feeling attached to a perceptual state because there are enough differences, between face-to-face seeing and picture seeing, in perceptual cues visual FOP relies on, so that we should not invoke further cognitive states in order to establish their difference concerning FOP. This chapter is important, insofar as visual FOP is neglected in both neuroscience and philosophy.

Issues in (§1.3.3, chapter 4) and (§1.3.4, chapter 5) develop a comprehensive, under-researched, new view of picture perception, in relation to vision-for-action.

2 Through the Forest of Motor Representations

The importance of investigating the mind starting from the naturalization of the mental antecedents of actions has been recently and excellently brought to the attention of the philosophy of mind (Nanay 2013b). Accordingly, following neuroscience, and using different labels, several philosophers have addressed the idea of the presence of a single representational mechanism lying in between (visual) perceptual processes and motor processes involved in different functions and useful for shaping suitable action performances: a MR[1]. MRs are the naturalized mental antecedents of action – I am excluding mental action here.

This chapter presents a new, non-monolithic view of MRs, according to which, contrarily to the received view, when looking at in between (visual) *perceptual processes* and *motor processes*, we find not only a single *representational mechanism* with different functions, but an ensemble of different sub-representational phenomena, each of which with a different function. This new view is able to avoid several issues emerging from the literature and to address something the literature is silent about, which however turns out to be crucial for a theory of MRs. Before developing my account, I need to sketch the basic positions in the literature about MRs.

2.1 The Positions in Play

In order to sketch the basic positions held in the literature, a premise on neurophysiology is needed.

As said, a common ground to link visual perception with action is the TVSM, which suggests the presence, in humans and other mammals, of a separation of two main visual pathways, grounded on distinct anatomo-functional structures (Milner and Goodale 1995): one for visual recognition, the ventral stream, and one for visually guided action, the dorsal stream. They can be dissociated due to cortical lesions. Lesions in the dorsal stream (the occipito-parietal network from the primary visual cortex to the posterior parietal cortex) impair one's ability to use what one sees to guide action (optic ataxia), but not object recognition; lesions in the ventral stream (the occipito-temporal network from the primary visual cortex to the infertemporal cortex) impair one's ability to recognize things in the visual world (visual agnosia), but not action guidance (Jacob and Jeannerod 2003). Moreover, we have behavioral studies of normal subjects involving visual illusions that can deceive the ventral stream but not the dorsal one; thus, it seems that, unlike ventral perception, dorsal perception is completely inaccessible to consciousness (I'll come back to this in § 2.4.2).

Starting from this specification, we now come to the list of the positions about MRs. I should specify that I will just mention the positions in the literature which propose a naturalized account of these mental antecedents of action, namely, those that follow neuroscience in order to build an idea of MRs. Indeed, I agree that, in order to understand MRs, we have to turn to empirical science (Nanay 2014a), something not always pursued in the literature about action-representations.

---

[1] I will ignore the different labels found in literature and simply talk about MRs.

2.1.1 MRs are Dorsal Phenomena, not Consciously Accessible and Represent Action Goals, Bodily Movements and Action Properties

A widely agreed idea about MRs is that they are due to the dorsal stream (Pacherie 2000, 2011) and, given its encapsulation, MRs are not normally consciously accessible (2000: Sec. 5; 2007: 8; 2011: 14; 2006 14; 2002: 63; Jacob and Jeannerod 2003: 252–255, 186, and Sec. 6.4; see Ch. 6; see Brogaard 2011: 1094; see Jeannerod 1994, 1997). MRs might become conscious by accessing the short-term memory long enough, when action is blocked, delayed, or through top-down attentional amplification (Jeannerod 1994; Pacherie 2006: 8; 2008: 195, 198; 2000: Sec. 5; p. 104). Also, since dorsal perception is not sensitive to certain perceptual illusions, thus MRs cannot therefore access ventral conscious semantic representations (I'll suggest this is not completely true in § 2.4.2) – as suggested by cases of visual agnosia (Pacherie 2008: 196; 2000: 411-412). However, sometimes a slight inter-streams interaction is allowed (following Jacob and Jeannerod 2003) (I'll come back to this in § 2.4.1) and it is suggested that, while MRs are dorsal processes, pragmatic and semantic representations should not rely on the ventral/dorsal anatomical distinction (Pacherie 2000: 411, 412). Moreover, MRs guide action as long as it unfolds (Pacherie 2002: 61; 2010: 10, 11; 2011: 14 and sec. 4; 2006: 14; see also Nanay 2014a: 4). In general, MRs represent an action goal (Jacob and Jeannerod 2003; Pacherie 2000: 409; 2007: 2; 2011:13; 2002: 71-72), which determines the type of grip chosen for motor interaction (Pacherie 2008: 186; 2006: 8; 2011: 10, 13; 2007: 2, 8; 2000: 409, Sec. 5; 2011:13; 2002: 71-72; Jacob 2005), and which is computed in egocentric coordinates (Pacherie 2011: 10; 2002: 70,71; 2000: 413; see Jacob and Jeannerod 2003: Ch. 8.2; see §§ 2.3.1, 2.3.2, 2.4.4) – the literature on MRs privileges the case of grasping, insofar as this specific motor act is the most studied one in the neuroscientific literature about visuomotor behavior; while I will follow the literature in privileging this kind of motor act, it is worth noting that the discussion can be extended to most of the motor acts we are able to perform (see Rietveld and Kiverstein 2014). Furthermore, MR functions 'fall between' a sensory function (extracting objects' features relevant for action) and a motor one (encoding motor acts) (see § 2.3.1) and those two aspects of the content of MRs (goal and object features) are not separate components of the content (see §§ 2.3.1, 2.3.2, 2.4.4). Finally, MRs also represent how our body has to move in a given situation, being grounded on the simulation of the required motor acts (Jacob and Jeannerod 2003: Ch. 6; Jeannerod 2006). This idea about the functions of MRs is also shared by Gallese (2000, 2005, 2007, 2009). However, it is not clear how MRs can select a singular motor act with respect to different action possibilities that a singular object can offer (this problem is recognized by Pacherie 2002: 71-72) (I provide an answer in §§ 2.4.3, 2.4.4).

2.1.2 MRs are subserved by both Streams, can be Conscious and Represent only Action Properties, not Goals or Bodily Movements

Another view is that MRs represent action properties extracted from the object's geometrical properties relevant/necessary for action performance (Nanay 2011c, 2013 b: 39; 2014a: 4; 2013 a: 75). However, while for Nanay (2011c) MRs also encode action goals, for Nanay (2013 b: 41, 42) this is not the case and they do not represent bodily movements (Nanay 2013 b: 41). Furthermore, while dorsal perception plays an important role in the implementation of MRs, it is not the only one (Nanay 2013a: 1055, 2013b: 3.4, p. 64). To my knowledge, Nanay is the only one to explicitly suggest that MRs are subserved by both streams and can be conscious, although they are typically unconscious. This idea seems to follow from empirical evidence about a lack of a clear-cut distinction between the two streams' processing, about the possibility of dorsal "consciousness", the multimodality of dorsal perception and the sensitivity of MRs to top-down factors like the subject's affective life, her/his language skills (Morgado et al. 2011; Pulvermüller and Hauk 2005; Stein et al. 2004; Gentilucci et al. 1995; for a specific discussion of these empirical results see Nanay 2013 b: 3.4). However, dorsal processing is supposed to be quick and automatic and thus should not be sensitive to top-down influences. Therefore, MRs are not fully exhausted by dorsal perception. Finally, it is suggested that the expression "dorsal vision" deals with anatomical criteria, while action-guiding vision deals with functional criteria and we should not use anatomical data in analyzing the cognitive impenetrability claim (Nanay 2013 a: 1058) confusing the functional level (vision-for-action) and the anatomical level (dorsal stream) (p. 1055).

2.1.3 MRs arise from the Activity of the Premotor Cortex and mainly encode Goals

Finally, there is the view according to which MR processing lies in the activity of premotor areas, without clarifying the precise cortical circuits, nor addressing the ventral/dorsal issue (Butterfill and Sinigaglia 2012). Following this view, the main function of MRs is to represent action outcomes. However, beside this main function, they can also represent ways of acting, objects (properties) on which actions are performed (2012: 137) - this is because action outcomes often specify both a way of acting and also what to act on. Also, they are useful in planning and in monitoring (p. 123, 124; see also Gallese and Sinigaglia 2011; Pacherie 2000: 410–3; Gallese 2000). I'll discuss this account more extensively later (§ 2.3.2)[2].

2.1.4 Philosophical Problems with (the literature on) MRs

Reading this list, several philosophical issues remain unspecified for a theory of MRs: (1) There are accounts for which the same MR has not the same function (e.g. for some commentators the function is to represent the goal, for others it represents motor commands, and for others still it is both). We should precisely address all MR functions; (2) more than just one function is addressed for a single MR, which can be engaged in different tasks. The doubt is whether and how the same representation can deal with so many functions, enslaving different tasks, insofar as every representation should have a singular content, a singular vehicle, and thus, a singular function (what the content is, in this case, what the function is about: goals, action properties, etc.), being this aspect crucial in determining the accuracy conditions of the representation (this is widely agreed in the philosophical literature about perception and neuroscience, see Siegel 2006, 2010a, b; Nanay 2010c; Jacob and Jeannerod 2003); (3) There is no clear-cut distinction between the functions addressed in the literature, it being unclear when one ends and the other begins. This issue is closely related to the following point: (4) While MRs are taken to be mainly dorsal phenomena, it is not clear as to how those functions are precisely subserved by the dorsal stream: even confining them to the activity of the dorsal stream, we do not know the precise dorsal neural correlates these functions arise from. Another issue (5) is that MRs are mainly, but not totally, dorsal phenomena. Ventral processing is involved in MR activity, in the light of evidence about the interplay between the two streams. If the precise dorsal contribution has failed to be addressed, the precise nature of the ventral contribution – and the possibility of conscious accessibility - is even more

---

[2] Here I do not talk about the direction of fit, the format and the link with imagery of MRs.

neglected. Finally, (6) there is no reference to any emotional component for MRs. However, neuroscientific results strongly suggest the presence of a further affective-emotional component of MRs.

2.2 A new perspective on MRs

The advantage of my proposal is that, maintaining the idea that MRs are mainly dorsal phenomena, but depend on ventral processing as well, I develop a theory of MRs able to avoid the aforementioned issues (§ 2.1.4), by clarifying what are the functions and what is the precise geography of the neural underpinnings of MRs. Indeed, contrarily to what is generally proposed, I argue that a singular MR can be decomposed into different motor representational *sub-components*[3] with different functions (this resolves issues 1 and 2 mentioned in § 2.1.4). For each of these components, I address the precise function and the precise neural underpinning in relation to both streams (this resolves issues 4 and 5), showing that these components are deeply interconnected, but in principle discernible (and this resolves issue 3). I also add an emotional component for MRs (resolving issue 6).

Summarizing, my aim is twofold: to defend a non-monolithic view of MRs and to suggest, on the basis of this view, the precise relation they entertain with the two visual streams. Here's the way I will develop my theory.

Since MRs are mainly dorsal phenomena, in the first part of the chapter, I develop my decomposing strategy (§ 2.3), focusing on the dorsal contribution to MRs and suggesting that MRs can be decomposed into an ensemble of different sub-representations, each one subserving a singular function, necessary for MR processing: (§ 2.3.1) the visuomotor component and the simulative one; the goal-related component (§ 2.3.2); the egocentric component (§ 2.3.3).

However, MRs also depend on ventral perception. Thus, in the second part of the chapter, I suggest the importance of ventral processing for the functions addressed above which mainly pertain to dorsal perception, showing that MRs depend on the dorsal/ventral interplay (§ 2.4). Then, I propose an unexpected function (§ 2.5) and an emotional component for MRs (§ 2.6), which are completely neglected in the literature.

Constructing my account is important for two reasons. Too often, MRs are invoked in terms of the perception of action of others, at the expense of the perception of action for the acting subject (Cook et al. 2014) – (see § 2.3.2). Also, though evidence about visuomotor behavior is constantly accumulating, we still lack a philosophical theory able to support it. My theory reconciles philosophy with neuroscience. So, let's start with the decomposing strategy.

2.3 The Decomposing Strategy

This first part of the chapter limits the discussion to the neurological geography of the dorsal stream (and its projections), leaving aside the ventral contributions.

Saying that the functions of MRs primarily rely on dorsal activity is not trivial and, for this reason, may be too general a claim. Indeed, the dorsal stream is an extremely complex pathway whose activity starts from the primary visual cortex, V1, and, passing

---

[3] Important specification: for Pacherie different action representations deal with different stages of action specification (2008: 195; 2002: 67), concerning a vertical differentiation about intentions. Differently, I argue that the same floor of action, that is MRs, can be decomposed into different sub-representations: I am drawing a horizontal, not vertical decomposition. This pertains to the lowest level of motor intentions and the linked level of MRs.

through the parietal lobe, projects to the premotor areas (with some projections until the motor one). Nevertheless, the complex connectivity of its cortical chunks should not lead us to couple MRs with dorsal activity, without specifying the nature of this coupling. Indeed, when we focus on a deeper level of grain, we find that different dorsal (sub)pathways subserve different representational functions. The sum of these processes shapes, at the high level of grain, the (apparent) single phenomenon we call MR. If we do not properly draw a distinction between those different processes, we risk using the word "MR" – with respect to those functions - as a folk notion: it makes no sense to talk about a general MR we can point our finger at when we locate it on the dorsal stream, just as we cannot generally point our finger at our belly and say: "well, this is where you find digestion". In other words, an MR is the ensemble of different processes whose complex organization gives us the impression of a single phenomenon; but this is only an impression. Ignoring this leads us to fall into the problems mentioned in (§ 2.1.4). I would like to make my argument transparent here by pointing out that the separability of the functions of the sub-representations composing an MR are individuated through the separability of the anatomo-functional portions of our visuomotor system. That is, my argument is based on separability of functions in neural areas (see also Grafton and Hamilton 2007; Thill et al. 2013; cfr. with footnote 3).

Note that the dorsal stream, which is involved in several important visuomotor tasks, is divided (at least) into two: the dorso-dorsal stream (D-D, following the division of the intraparietal sulcus, which subdivides the posterior parietal lobe, it is related to the superior parietal lobule (SPL) - also known as the dorso-medial circuit, which projects to the dorsal premotor cortex) and the ventro-dorsal stream (V-D, related to the inferior parietal lobule, (IPL) - also known as the dorso-lateral circuit, which projects to the ventral premotor cortex) (Gallese 2007; Turella and Lignau 2014; Rizzolatti and Matelli 2003). Though both chunks of the dorsal stream are crucial for motor interaction (especially grasping), V-D and the ensemble of its parietal-premotor networks (AIP-F5 and VIP-F4), is the crucial pathway, both in humans and non-human primates, concerning the representational components of MRs listed above in (§ 2.1) - with respect to D-D, mostly involved in proprioceptive input, but with an important visual contribution (Gallese 2007):

(a) The Visuomotor Component and the Simulative one
(b) The one linked with Goal/Outcomes
(c) The Egocentric Component

I will now propose a neurological guided tour of the V-D pausing in each cortical site dealing with an MR component.

For the sake of coherence toward neuroscience, I should specify that, of course, while my analysis is more accurate than those in the philosophical literature, it is not so technical as those in the neurophysiological literature. But this is not a neuroscientific review of the neural underpinnings of action-guiding vision (Chinellato and del Pobil 2015). It is a philosophical analysis that considers the necessary empirical background in order to avoid a too simplicistic monolithic view of MRs.

Now the decomposition, which starts from the visuomotor component, accompanied by the simulatiove one.

2.3.1 The Visuomotor Component and the Simulative one

Objects exhibit geometrical properties (e.g. size, shape, texture) that are, from a motor point of view, action/motor properties, to the extent that they permit a precise action possibility[4] satisfiable with a precise motor act. For example the geometrical features of a mug are action properties permitting an action possibility (grasping), which can be satisfied by a proper motor act: a power grip. This important function of MRs is due to their visuomotor component. This visuomotor transformation relies on a well defined parietal-premotor network lying in between the posterior parietal cortex and the ventral premotor cortex, that is, a precise portion of the V-D, whose main components for these tasks are the anterior intraparietal (AIP) area and area F5, in the most rostral part of the ventral premotor cortex (for a review of the leading role of AIP-F5 in the detection of action possibilities and the related visuomotor transformation of object properties into action properties and then of action properties into motor acts see Borghi and Riggio 2015, Raos et al. 2006, Romero et al. 2014, Castiello 2005; Castiello and Begliomini 2008 Chinellato and del Pobil 2015; Kandel 2013: Chap. 19; Janssen and Scherberger 2015; concerning lesion studies see Andersen et al. 2014; Turella and Lignau 2014; Graziano 2009) – see footnote 19. AIP is one of the end-stage areas of the dorsal stream. AIP neurons respond selectively to objects during both passive fixation and grasping, extracting visual object information concerning action possibilities for grasping purposes (for a review see Romero et al. 2014, Raos et al. 2006); then, they relay this information to F5 neurons, with which AIP is directly connected (Borra et al. 2008), which then activate the primary motor cortex. In F5 we find visuomotor canonical neurons, which use the information received by AIP about action properties of the objects and compute the motor commands in order to interact with them. Also canonical neurons respond during object fixation, regardless the actual execution of an action. In canonical neurons activity, there is a strict congruence between their high selectivity for a particular type of prehension (executed grip) and the visual selectivity for objects that, although differing in shape, require prehension in order to be grasped (for a review see Raos et al. 2006). Imagine you have to grasp first a little box that can be contained inside your hand, and then a little stone: although their shape is different, these objects show the same action property; therefore the motor acts satisfying this action possibility (grasping the object with the whole hand) are the same. This is a first example of how features of objects are read (during visuomotor transformation) as contents of a (sensori)motor nature. Representational presence concerning the visuomotor component is also suggested by an automatic process of grip formation taking place during the transportation of the hand in which the fingers are preshaped long before the hand touches the object. At about 60% of its transportation the hand reaches its widest opening, or *maximum grip aperture* (hence forth: MGA) and the size of the finger-grip at MGA (though much larger than the object to be grasped) is linearly correlated with the size of the object: this process is largely anticipatory and pertains to an automatic action representation, not to a mere on-line adaptation of the motor commands to the object (Jeannerod 2006: 5; for the shown

---

[4] There is a crucial point here. The concept of action/motor property/possibility recalls the famous notion of affordance, proposed by Gibson (1979), which has captured the interest of both neuroscientists and philosophers in the last twenty years (Borghi and Riggio 2015; Thill et al. 2013; Ferretti, forthcoming; Jacob and Jeannerod 2013; Nanay 2013b; Chemero 2009). Affordances involve both perception and action, insofar as they consist in the visual perception of the invitation to action that is offered to the active subject by the objects she/he deals with in the environment. However, the notion of affordance is strictly related to the gibsonian idea of a direct visual perception of action possibilities, to the extent that we do not need to use any visual representation in order to detect them (for a complete analysis see Jacob and Jeannerod 2013). However, despite the original gibsonian anti-representational view of affordances (for an analysis see Chemero, 2009, Nanay 2013b), a lot of neuroscientists and philosophers agree that action possibilities are perceived through perceptual representations (Nanay 2013b, Jacob and Jeannerod 2003; for an argument built on empirical evidence see Ferretti, forthcoming) and avoid the use of the term "affordance" with an anti-representationalist stance. Here I follow this second view insofar as my account is framed in terms of MRs. Hence, I avoid the use of the term "affordance" and refer to the more neutral expression "action property/possibility".

automaticity of the process see Gentilucci et al. 1995; for a review see Nowak and Hermsdörfer 2009). These *visuomotor phenomena* are the crucial process of the visuomotor component[5].

A further specification. These cortical circuits exhibit both the visuomotor transformation processes and a mechanism of visuomotor resonance during object fixation, regardless of the action execution: motor simulation. Simulation is an automatic mechanism with a perceptual function to facilitate the motor preparation (Gallese 2009)[6]. Motor activation frames the represented action within the constraints of a real action[7]: represented actions correspond to covert actions as a neurophysiological simulation of the mechanisms normally involved in the physical action generation" (Jeannerod 2006: 130-131)[8]. However, we have an issue here: for Jeannerod the representation of the action is basically the simulation of the action, even though he admits that the representations for executing and simulating do not completely overlap, allowing this distinction even in the absence of sensory reafferences" (Jeannerod 2006: 131).

In spite of this incomplete overlap I would like to suggest a distinction between motor simulation and visuomotor representation. Simulation concerns motor response during fixation with respect to the activity of both AIP and F5 visuomotor canonical neurons. The visuomotor representational mechanism concerns the interplay of (the resonance of both) AIP and F5 (the former resonating in the encoding of layout properties of objects as action properties and the latter in encoding the motor act with respect to these action properties) for the visuomotor transformation of object properties (both without the necessity of action performance). Therefore, I maintain that motor simulation and MRs are deeply linked (Jeannerod 2006: 130-131), but reformulate Jeannerod's idea that the (overt) execution of an action is necessarily preceded by its (covert) simulation, while a (covert) simulation is not necessarily followed by an (overt) execution of that action. The simulation is possible only given the result of the visuomotor transformation.

All I said here recalls – and is deeply in line with - the idea of the common coding theory of perception and action planning, according to which perceptual contents and motor programs which instantiate action plans are coded in a common representational process to the extent that seeing an object activates the action associated with that object. In this view, perceptual representations and MRs are linked by shared representations in

---

[5] However, AIP needs the help of F5 for encoding action properties; thus, the encoding of motor acts and action properties cannot be properly divided (Romero et al. 2014, Theys et al. 2015). This is a form of motor perception (Fadiga et al. 2000: 165, 176; Jacob and Jeannerod 2003: 177). Accordingly, there is also evidence that in the AIP-F5 circuit 3D-shape selective neurons are co-localized with neurons showing motor-related activity a close relation of visual and motor information on the same clusters of neurons (Theys et al. 2015; Romero et al. 2013). This is in line with the evidence of several action-modulated perceptual effects that improve processing of action relevant visual features (Gutteling et al. 2015) and with the evidence that action preparation originating from activation in higher-order visuomotor control brain areas downstream from early visual cortex in the dorsal stream modulates visual processing according to the visuomotor constraints of the action that is encoded (van Elk et al. 2010; Gutteling et al. 2013).

[6] "What is a 'plan' to act? It is a simulated potential action" (Gallese 2007: 7; see also Gallese 2000; Jeannerod 2006; for a review see Borghi and Cimatti 2010; Decety and Grezes 2006). Motor simulation is a mental rehearsal of the motor acts able to satisfy an action possibility related to the object (Tipper et al. 2006).

[7] …though neural commands for muscular contractions are effectively present, but simultaneously blocked by inhibitory mechanisms (Jeannerod 2006: 2.3.3).

[8] Overt action execution is necessarily preceded by its covert representation and simulation. Covert representation and simulation are not necessarily followed by overt execution. Representation can be detached from execution, existing on its own (Jeannerod 2006: 2; chap. 2, 6). Also, motor activation is highly specific to the action that is represented (Id.), involving the main neural structures needed in action execution.

a common code for both perception and action, insofar as actions are coded in terms of the perceivable effects they can generate (Prinz 1984, 1987, 1990, 1997; Hommel et al., 2001; for a complete review see Prinz et al. 2013; for an analysis of the compatibility between these sets of evidence and the common coding theory see Borghi and Riggio 2015: 7).

Summarizing, when I am looking at the cup of coffee on my desk, V-D responds to those 3D geometrical properties of objects that serve such visuomotor tasks as grasping them. The AIP-F5 parietal-premotor network is crucial in translating those geometrical features into action properties and then into motor acts. Thus we can grasp the cup. First of all, AIP detects the geometrical features of the handle that exhibit precise motor characteristics. This means that shape, texture, size are encoded. Thus, the geometrical features are read as action properties. This information is sent to F5, which, given the information received by the AIP, computes the most suitable motor act (with respect to my motor repertoire, say, a power grip) in order to grasp the handle of the cup. At the same time, during this translation, the simulation of the appropriate motor behavior is encoded: the visuomotor brain is both perceiving the action possibility and simulating its readiness for potential related motor interaction, setting a motor act "in the quiver"[9]. I'll refer to this particular component and the brain areas it depends on several times in this work. I'll be back to this in (§ 3), in particular in (§ 3.1) and in (§ 4), in particular in (§ 4.3).

2.3.2 Goals/Outcomes

Action performance presupposes a some sort of representation of the goal (for a review see Nanay 2011c, 2014a, for critics see Nanay 2013b) and this is due to the second of MRs component that deals with goals/outcomes. Butterfill and Sinigaglia (2012) have recently described the goal-related nature of MRs. As they suggest, "the representation of the goal is not a mere representation of a sequence of joint displacements, insofar as it captures something shared by different patterns of joint displacements/postures involving different effectors - e.g. hand, mouth, use of normal pliers (grasping requires closing the hand), use of reverse pliers (grasping requires opening the hand) (p. 121) - and discerns between the same sequence of joint displacements in different contexts, depending on distal outcomes, e.g. eating or throwing the object after having grasped it". Indeed, the joint displacements realizing grasping in one context might in another situation realize scratching (p. 122)". This is confirmed by evidence that "markers of motor processing, such as a pattern of neuronal discharge or motor-evoked potentials, carry information about action outcomes" (p. 122) and not joint displacements. On this matter, the Authors report experiments with varying kinematic features while holding constant the outcome achieved using different effectors (Rizzolatti et al. 1988, 2001; Cattaneo et al. 2010), studies where the same action outcome requires closing or opening the hand depending on the tool used (Umiltà et al. 2008; Cattaneo et al. 2009; Rochat et al. 2010), experiments with varying action outcome while holding kinematic constant (Fogassi et al. 2005; Bonini et al. 2010; Cattaneo et al. 2007), and studies about the same grasping movements performed in the presence/absence of a target object (Umiltà et al. 2001; Villiger et al. 2010) or in the presence of objects which could, or could not, be grasped with such movements (Koch et al. 2010; see Butterfill and Sinigaglia 2012: 121)[10] - (for the empirical reference see Butterfill and Sinigaglia 2012).

---

[9] For the behavioral counterpart of these experiments see (Borghi and Riggio 2015).

[10] They also suggest that MRs are useful in motor planning and monitoring. Since the literature contains a vast amount of evidence I can avoid this point (Butterfill and Sinigaglia 2012: 123, 124; Nanay 2014a: 4;

While the goal-related component can be completely exhausted by Butterfill and Sinigaglia's account, there is an important point concerning the evidence exposed above, which I need to address in order to build my theory.

First of all, note that area F5 (§ 2.3.1) contains (at least) two large groups of neurons: the first is that of (A) purely motor neurons, whose activation is exclusively connected to actual movements. They constitute the overall majority of all F5 neurons, and belong to two kinds: (A1) neurons that fire whenever a movement is performed, and (A2) goal-related neurons, which code only the achievement of a goal regardless of the effector (i.e. the particular limb employed) (Fadiga et al. 2000, Rizzolatti and Matelli 2003). But we also have (B) visuomotor neurons, also distinguished into two groups: (B1) canonical neurons described in (§ 2.3.1); (B2) mirror neurons, which respond when the monkey – see footnote 19 - observes an action performed by another individual, or when it performs the same or a similar action (Cook et al. 2014).

The experiments reported by Butterfill and Sinigaglia – I mentioned above - pertain to different families of neurons. However, most of them are about mirror neurons (B2) (Rizzolatti et al. 1988; Cattaneo et al. 2010; Cattaneo et al. 2009; Fogassi et al. 2005; Koch et al. 2010; Umiltà et al. 2001; Villiger et al. 2010). Since here I am interested in the individual dimension of our motor behavior, I would like not to focus on the evidence about the representation of goals during observation (B2) (for a complete review of these very important results see Sinigaglia and Butterfill 2015a, b; Gallese and Sinigaglia 2011b, 2013, 2014; Sinigaglia 2013; Rizzolatti and Sinigaglia 2008, 2013) - even if it is very useful, as shown by Butterfill and Sinigaglia - for the purpose of building my theory of MRs and explaining the representations allowing us to act. However, if we endorse my suggestion, there are important experiments quoted in their account, namely, that of Umiltà et al. (2008) - about grasping with normal and inverse pliers - which is the same experiment as Cattaneo et al. (2009), but without the mirror counterpart – that of Bonini et al. (2010) – which, however, is about both action organization and understanding concerning the activity of the ventral premotor and inferior parietal cortices – and that of Cattaneo et al. (2007) - concerning how parietal and premotor neurons involved in the encoding of a very specific motor act (e.g., grasping) show a significant different activation when the motor act is part of actions linked to different goals (e.g., grasping for eating vs. grasping for placing), see above (for the empirical reference see Butterfill and Sinigaglia 2012).

So, I would like to point out here that into enquiring the goal-relatedness - of the dorsal underpinnings - of MRs, my research focuses on those last mentioned kinds of experiments on the individual side. For example among the neurons involved in representing outcomes are those called goal-related neurons (A2), which don't encode elementary movements as joint displacements, but motor acts (coordinated movements with specific purposes) (Rizzolatti et al. 1988). They are interesting because the same elementary movement activating a neuron during a specific motor act (e.g., grasping) doesn't activate it during a different motor act (e.g., scratching). There are thus different groups of neurons from F5: grasping neurons, grasping-with-the-mouth neurons, hugging neurons, etc. This is possible because during our ontogenetic development, the pruning of our neural networks under the pressure of experience selects in F5 the neural populations linked to the (representations of the) most effective motor acts. This learning mechanism is called "motor reinforcement".

This is important because F5 is a *vocabulary* whose words are neural populations representing one kind of *motor act* as the ensemble of different motor words (rather than

---

Pacherie 2008: 189, 90; 2002: 61; 2010: 10, 11; 2011: 14 and sec. 4; 2006:14; for a more general discussion about action control see Wong 2009, 2010, 2015: Shepherd 2014, 2015a, b). I am primarily interested in the MR functions necessary to transform the sensory input into motor output in order to explain the mental antecedents of actions.

a simple *movement*). The referent of these "words" can be of different generality: the goal, the execution, etc. (Rizzolatti *et al.* 1988; Gallese and Metzinger 2003, 367). As Jeannerod puts it, in the 'motor vocabulary' actions are encoded element by element (2006: 12). Due to its somatotopic organization, this vocabulary provides computational efficiency, insofar as the appearance of the graspable object in the visual space will–immediately retrieve motor words with the description codifying the appropriate motor act (Gallese and Metzinger 2003, 367-368)[11], recorded in the motor vocabulary as an internal copy of an action (Fadiga et al. 2000). This is closely related with the activity of the visuomotor component (§ 2.3.1). Of course, goal encoding is not exhausted by these phenomena, insofar as the motor system plays a crucial role in the goal-component of MRs (Jacob and Jeannerod 2003; Jeannerod 2006). However, I am particularly interested in the dorsal contributions – which are at the heart of the debate - leaving aside its projections to the motor system.

So, even if I completely endorse the philosophical analysis by Butterfill and Sinigaglia, I think that in order to properly investigate the goal-related component we should report only the evidence concerning the individual aspect of the goal-relatedness of these neural underpinnings, and not the mirror-social aspect.

An important related point: for Jeannerod actions are represented in terms of their goal, but the goal is only part of the content of the action representation[12]. By representing the goal, we can answer the question of 'What the action is about', but not the question of 'How to do it', insofar as this question requires motor simulation to be answered (Jeannerod 2006: 134). I completely agree with Jeannerod that actions are represented in terms of their goal, which are only part of the content of the MR. However, I think that goal representation is also concerned with the "how to do it" question. Computing a motor command depends on the goal to reach and, *vice versa*, trying to achieve a goal depends on the possibility of performing a motor act among those skills one has. The details of this interplay become clearer in (§§ 2.4.3, 2.4.4).

Summing up, visuomotor transformation and motor simulation are necessarily accompanied by the goal we have in mind (e.g. grasping the cup in order to lift it), on the basis of which we compute the motor act (e.g. a precision grip permitting me to grasp the cup and lift it to bring it to my mouth). All I said is strengthened by evidence that during action planning the action goal dominates over the hand grip (van Elk et al. 2011) and specific motor programs are selected on the basis of the action outcome (Bonini et al. 2012; Borghi and Riggio 2015: 7). However, we are usually faced with more than one motor possibility, and we have to figure out, with respect to the goal, which is the most appropriate way to interact (to grasp the cup with a power or a precision grip to better transport it to the mouth); matters on this point will become clearer in (§§ 2.3.3, 2.4.3, 2.4.4). The next section is about the egocentric component.

2.3.3 The Egocentric/Peripersonal Component

When we try to reach and grasp an object, we need to represent where the object is located with respect to us[13], since we can actively interact only with those objects presented within our peripersonal-action space. Visuomotor representation and motor

---

[11] Different populations of goal-related neurons can be described as being different representations. I avoid this point here.

[12] This is in line with my idea that different portions of the information encoded by MRs are subserved by different sub-representations.

[13] For a philosophical defense see (Nanay 2011); for criticism see (Nanay 2013 b: 41-42).

simulation are deeply dependent on the peripersonal spatial location of the object (Costantini et al. 2010, 2011a, b; Holmes and Spence 2004; ter Horst et al. 2011; Borghi and Riggio 2015; Turella and Lignau 2014).

This is due to corto-cortical functional interconnections of different areas within the circuits of grasping (Castiello 2005, Turella and Lignau 2014): namely, the two main – but not the only, see footnote 19 - parietal-premotor circuits in the ventro-dorsal stream: AIP-F5 and VIP-F4 (Rizzolatti and Matelli 2003; Gallese 2007; Chinellato and Del Pobil 2015). The circuit crucially involved in the mapping of our peripersonal space is composed of the ventral intraparietal area (VIP), located in the fundus of the intraparietal sulcus, and area F4, which occupies the posterior sector of the ventral premotor cortex, next to F5, with which it is adjacent (see Gallese 2007, Rizzolatti Matelli 2003). VIP neurons respond to both visual and tactile stimuli and the visual receptive field encompasses a three-dimensional spatial region around the tactile receptive field (peripersonal space). As for VIP neurons, most of the receptive fields of the F4 neurons do not change position with respect to the observer when the eyes move: F4 response does not signal retinal positions, but positions in space relative to the observer with respect to different body parts in egocentric coordinates and not to a single reference point. Multiple representations linked to different, coordinated, effector-specific frames of reference (e.g. head-centered, torso-centered) are present (Graziano 2009; Pesaran 2006 for the relation of these egocentric representations with the goal of action). F4 activity is due to simulated motor action directed towards a particular spatial location, which, in turn, creates a motor space. When a visual stimulus is presented, it directly evokes the simulation of the congruent motor schema which, regardless of the execution, maps the stimulus position in motor terms (Rizzolatti and Matelli 2003, Gallese 2007, for a philosophical analysis of space in relation to the TVSM see Briscoe 2009).

Interestingly for this chapter, the AIP-F5 and VIP-F4 circuits are extremely interconnected from an anatomo-functional point of view: largely segregated parietofrontal connections link the rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4) (Luppino et al. 1999). Indeed, the VIP- F4 circuit codes the peripersonal space, which coincides with the motor space for arm reaching in which motor interaction is encoded by AIP-F5 (Borghi and Riggio 2015: 9). For example, F5 canonical neurons respond only to those objects presented in the peripersonal space (Bonini et al. 2014). This is due to the connections between area F5 and area F4 (Borghi and Riggio 2015: 7; Costantini et al. 2011a; Bonini et al., 2014). This spatial constraint holds in general for the ventral premotor cortex as well (Turella and Lignau 2014) – see footnote 19. All this means that the visuomotor transformation (performed by AIP-F5) is bound to the peripersonal/egocentric action space (performed by VIP-F4) – but see footnote 19. I'll be back to this particular component and the brain areas it depends on in (§ 4), in particular in (§ 4.5).

Returning to the example of the cup, I can perform an appropriate motor act upon its handle because I can represent it as affording an action, due to the interplay of the visuomotor (and simulative) and goal related components. However, representing the cup as affording an action, means to represent the cup as reachable for me (that is, in my peripersonal space).

To conclude this first part of the chapter, this decomposing strategy has suggested ensembles of different sub-representations dealing with different functions subserved by different neural correlates of the V-D at the basis of MRs. This is important, insofar as these functions are often mistakenly attributed to a single general MR, and too generally to dorsal perception. This is a first way to avoid the confusions reported in (§ 2.1.4).

However, I take MRs to depend on ventral perception and on interstream interplay. The second part of the chapter explores this dependence with respect to the functions of MRs reported above and adds some features that can be explained only by referring to ventral perception.

2.4 MRs between Ventral and Dorsal Perception

My decomposing strategy defends a new non-monolithic view of MRs by addressing their different and specific representational functions in relation to dorsal perception. However, we know that any sophisticated visual behavior requires interstream collaboration (Schenk and McIntosh 2010; Kravitz et al. 2011, 2013: Box 3, p. 42; Bruno and Battaglini 2008; McIntosh and Schenk 2009; Briscoe 2009). The literature about MRs neglects these results[14]. By discussing them here, I defend the second important idea of this chapter: MRs also rely on ventral processing, which shapes important computational aspects of their components outlined in the decomposing strategy. I also report important insights on dorsal perception neglected even by those claiming that MRs are only dorsal phenomena.

2.4.1 The Dorsal/Ventral Interplay in Action

Neurophysiology of vision suggests there is no clear-cut functional distinction between the streams at various points in perceptual processing (Schenk and McIntosh 2010; Gallese et al. 1999): they integrate in early visual areas by feedbacks allowing each pathway to affect the other (Deco et al. 2004) and by sharing common early visual inputs (Schneider 1995), insofar as both connect with the frontal eye field, so that eye movements initiated by one stream might also impact the other (Kravitz et al. 2013). This interplay is at the basis of vision for action (for a complete review see Chinellato and del Pobil 2015), suggesting that MRs cannot be genuinely dorsal phenomena. Indeed, interplay is at the basis of attentional processes and visual memory in vision-for-action: visual short-term memory representation of tools automatically evokes action possibilities in dorsal perception, priming visual memory for action-appropriate objects (Adamo and Ferber 2009; see also Kitadono and Humphreys 2009). Accordingly, dorsal perception is not only involved in on-line visuomotor coordination, because motor responses to action properties during off-line visual processing are possible by using visual memory even after 700 ms. after the object has been removed from view (Derbyshire et al. 2006: 95). Thus, (at least some) detection of action possibilities is not only dorsal, but results from an interstream interaction for off-line visual processing. Accordingly, we know that dorsal perception encodes action possibilities (unconsciously), whereas ventral perception does this consciously (Young 2006), following object semantic categorization (Gallese 2007: 3). Therefore, during delayed actions, dorsal perception plans and maintains coarse action goals, but at the time of execution, motor programming re-recruits detailed visual representations from ventral perception, particularly for grasping (Singhal et al. 2013) – see my (§ 2.4.3)[15]. Finally,

---

[14] Nanay simply pointed out that MRs are not exclusively dorsal phenomena. Here I add a meticulous cortical geography of MR functions with respect to both streams.

[15] Also, motor planning with visual feedback at the onset of the movement is driven primarily by real-time visuomotor computations of dorsal perception, whereas grasping remembered objects without visual feedback is driven primarily by the perceptual memory representations of ventral perception (Prime and

while the motor planning/programming distinction is a boundary between ventral and dorsal influences on action (Milner and Goodale 2008), interestingly, some visual agnosic patients show impaired action planning (high-level selection of action) with preserved action programming (low level parameterization), suggesting that planning depends on a perceptually based judgement of the awkwardness of alternative movements involving ventral perception, disrupted in visual agnosia (Dijkerman et al. 2009). Accordingly, misreaching affects not only extra-foveal visual targets, but also proprioceptive targets displaced from the direction of the gaze, suggesting that optic ataxia is not only a visuomotor disorder, and thus deep interstream collaboration in action (Jackson et al. 2009). Finally, both visual agnosia and optic ataxia only support dissociation between central visual processing and peripheral visual processing: ataxic patients cannot reach and grasp in peripheral vision, but can in central vision (Rossetti et al. 2003). This interstream interplay subserving MRs becomes clearer when talking about the deception, of vision-for-action, to illusions.

2.4.2 Dorsal Perception without Dorsal Deception?

A big issue about the dorsal/ventral dichotomy concerning vision-for-action is whether or not illusions can deceive it. While the original response with the 3-D version of the Ebbinghaus illusion was negative (Aglioti et al. 1995), results in a variety of different experimental settings showed that the Ebbinghaus/Titchener illusion similarly deceives both vision and grasping, by deceiving a common representation of object size used by both (Franz and Gegenfurtner 2008) – of course, that means that the representation is non-veridical, not really deceived; subjects or persons are who get deceived. Accordingly, we have a huge meta-analysis showing that the Müller-Lyer illusion (Bruno et al. 2008; Bruno and Franz 2009) influences grip-scaling, that it is similar on both visual perception and grasping when grasping is executed without visual feedback, while it reduces progressively as visual feedback is increasingly available (Franz et al. 2009) and similarly affects grip programming, even if errors can be corrected online if the hand is seen approaching the target; indeed, we have relative resistance of pointing movements to this illusion only if vision is available during action programming (Bruno et al. 2008; Smeets et al. 2002). Also, evidence that visuomotor reaction time is sensitive to both the Ponzo and Ebbinghaus–Titchener illusion suggested that vision-for-action can be subserved by ventral perception (Sperandio et al. 2009). While these results concern few illusions and few kinds of motor acts, a more general analysis revealed that visual perception seems to be fooled by illusions more than vision-for-action, only because dorsal perception is not attuned to contextual depth cues when action is automatic – e.g. MGA (Gentilucci et al. 1995), but when these cues are inaccurate, or usual dorsal sources of spatial information are unavailable, or the action is slow, visuomotor vision is misled by making use of outputs from ventral processing (I am rephrasing Briscoe 2009: 441). An important point: the real result by Aglioti et al. (1995) – which is the cornerstone of this debate - found that the Ebbinghaus illusion has a slighter effect on perception than on action showing only a greater sensitivity of ventral perception to illusions than dorsal (see Briscoe 2009: 436, 441).

Summing up, of course there are illusions incapable of deceiving action, but only visual perception. However, at least in the case of particular illusions, also vision-for-action can be deceived – something denied by Pacherie (cfr. with 2.4), because dorsal perception has ready task-specific access to sources of spatial content (contextual depth

---

Marotta 2013).

cues and other sources of 3-D spatial information) in the ventral stream. Thus, MRs can be deceived as well. This is in accordance with the idea that not all movements are mediated by dorsal perception. Ventral perception guides actions prior to their automatization (I'll come back to automaticity in § 2.4.4), and in this case, conscious perception is incorporated into dorsal perception (Milner and Goodale 2010: 83). Moreover, even those who deny cognitive access to dorsal perception[16] agree that conscious ventral vision affects dorsal vision-for-action (Brogaard 2011: 1094), insofar as illusions can affect the latter[17]. However, besides a general interstream interplay, precise ventral/ventro-dorsal connections shape MRs; the next section focuses on these.

2.4.3 The Connection between V-D and Ventral Perception

Everyday objects offer us a variety of action possibilities and thus different motor acts to perform upon them. The selection of the appropriate motor act does not rely only on the layout object properties (Cfr. with § 2.3.2), but also on what we intend to do with it, in relation to its functions and this is strictly related not only with what the object offers us, but also what we want to do with it. The interplay between the analysis of physical properties (pragmatic analysis) and object identity (semantic analysis) is due to connections lying between V-D and ventral perception and suggests the importance of ventral perception for MR functions. We saw in (§ 2.3.1) that AIP selects the geometrical properties to be translated into action properties and to be sent to F5 for the encoding of proper motor acts. Importantly, the action properties linked to the semantic functions of the object – e.g. think about the different grips we can use in order to use a pen in different ways: writing or placing the pen somewhere else - are possible because AIP receives information about the meaning of the object for high quality object recognition from both the inferior temporal (TEm, TE and TEO) and posterior parietal activity – thus, AIP is involved in object recognition (Fogassi and Luppino 2005: 627; Rizzolatti and Sinigaglia 2008; for a complete review, which I cannot offer here, of the complete functional processing of the pathways linking the early visual areas to the motor areas and about the relationships between the ventral stream and the ventro-dorsal stream see Chinellato and del Pobil 2015: 69; see also Theys et al. 2015). After the semantic analysis, the information processing from AIP to F5 results in a competition of the neural populations encoding different potential motor acts with respect to the action possibilities detected, on the basis of this semantic analysis – (cfr. with the last part of § 2.1.1) (Rizzolatti and Sinigaglia 2008: 36-38; Cisek 2007; Cisek and Kalaska 2010; Kandel et al. 2013: Chap. 19; van Elk et al. 2014; for a complete review see Thill et al. 2013)[18]. We have related evidence of a semantic component for motor processing (Helbig et al. 2006; Kalènine et al. 2013; Iachini et al. 2008; Tucker and Ellis 2004), and action preparation (Lindemann et al. 2006) given by the influence of ventral perception on action (Schenk and McIntosh 2010), as well a dorsal involvement in semantic encoding (Pulvermüller 2013: Box.1, Box.2) – we have also evidence that, in general, action usually shapes object categorization and that this influence is goal-related (Triberti et al. 2015). Also, MRs manage situations in which structural information and functional information

---

[16] There is no crucial evidence to suggest that dorsal processing cannot be conscious (Nanay 2015: 187), see (§§ 2.4.2, 2.4.5).

[17] I cannot survey here the complete literature concerning illusions and dissociation, for a more detailed account see (Bruno and Franz 2009; Bruno and Battaglini 2008; McIntosh and Schenk 2009; Bruno et al. 2008; Franz and Gegenfurtner 2008; Franz et al. 2009; Vishton et al. 2003).

[18] It is unclear whether all action possibilities are automatically activated and then some of them decay, or if a single action possibility is directly encoded among others. Borghi and Riggio (2015: 13, see also p. 8 for a discussion of the empirical results I cannot review here) go for the second hypothesis.

may conflict (Jax and Buxbaum 2010). This is in line with the evidence by McIntosh and Lashleya (2008) discussed by Nanay (2013b: 66) that top down factors influence our MRs – e.g. the brand of a matchbox influences our grip size when grasping it.

Summing up the visuomotor comonent of MRs is always supplemented by semantic object recognition through ventral perception. The most ventral chunk of the dorsal stream, V-D, with its projections to the ventral stream, is the cutting edge of dorsal perception involved in MRs functions, bringing together all the dorsal computational characteristics with an important ventral computational shade in relation to the goal-related component (§ 2.3.2) and the visuomotor one (§ 2.3.1). Thus, MRs rely on the ventral/dorsal interplay (cfr. with the account by Pacherie, Jacob and Jeannerod § 2.1.1). This also strengthens my remark about Jeannerod's idea that the encoding of goals and motor properties are interconnected (cfr. with §§ 2.3.2, 2.4.3, 2.4.4) and underlines the importance of goal encoding in order to compute action properties, both denied (Nanay 2013b) and endorsed (Nanay 2011c; Butterfill and Sinigaglia 2012) in the literature (cfr. with § 2.1). This specification is important because the dorsal stream contains another important chunk: the D-D. Pointing out the characteristics of the two chunks is crucial to understand the contribution of the different ventral and dorsal bifurcations, not of the visual brain in general, but of the dorsal stream in particular, to MRs.

2.4.4 The many facets of our Visuomotor Interactions

Here I want to specify that different specific MRs circuits, D-D and V-D[19], compute specific kinds of action possibilities. We can distinguish between variable and stable mechanisms detecting action possibilities. Stable action possibilities derive from stable/invariant properties of objects due to the associations between the visual aspects of a precise object and the motor response it produces that can be incorporated into an object memory-stored representation (e.g., we "know" that this object is graspable with a precision grip) (cfr. with § 2.3.2) (Borghi and Riggio 2015)[20]. Differently, variable action possibilities are linked to the actions we are about to perform, but deal with rather temporary object characteristics which are not memory-stored. V-D is responsible for the former, and by managing how our knowledge of objects influences the way we represent them. D-D is responsible for the latter, in the online interaction with objects (in new motor situations), encoding the orientation of the wrist for grasping the object under visual guidance, and continuously adjusting the grip to the object online. This is really important for the issue concerning the online processing of MRs – cfr. with footnote 10. Accordingly, MRs' detection of action possibilities is not so automatic, as agreed in the philosophical literature – indeed, the issue of automaticity of MRs is hotly

---

[19] In line with (§§ 2.4.1 2.4.2) V-D, D-D and ventral pathways are strictly interconnected and finish in cortical frontal areas (Borghi and Riggio 2015: 351). The same dependence between spatial encoding and the visuomotor one in the V-D, holds for the D-D. However, data – that I cannot review here - suggest the leading role of (both AIP-F5 and VIP-F4 in the) V-D concerning the visuomotor transformation (in the peripersonal space); for a critical review see (for a review see Castiello 2005, Castiello and Begliomini 2008; Turella and Lignau 2014; Kandel et al. 2013: 871; Janssen and Scherberger 2015; for the role played by the posterior parietal cortex MRs see Buneo ad Andersen 2006). All the evidence I reported is grounded on studies on both human and non-human primates and an important overlap between the different cortical areas is widely agreed (see Borghi and Riggio 2015: 3; Shikata 2003). This is not relevant for the purpose of this chapter.

[20] "Stable" does not mean that action possibilities are not processed and responded to online, but that they might need a certain degree of adjustment of the organism in relation to objects.

debated in the neuroscientific literature, in which it is suggested that the task ad the context are crucial in determining the "behavior" of the MR (see Borghi and Riggio 2015: 8; Girardi et al. 2010; van Elk et al. 2014; cfr. with my footnote 18). Also, MRs handle contexts where multiple objects are present, hence where multiple action possibilities are activated with respect to memory stored information (Pezzulo et al. 2010; Derbyshire et al. 2006) and the different semantic, functional, spatial, geometrical, context and task relations existing between objects, strongly affect motor responses (Borghi et al. 2012; Cisek 2007; Cisek and Kalaska 2010; Borghi and Riggio 2015). In line with this and the notion of MGA expressed in (§ 2.3.1), we know that while a macaque plans to grasp a single object, its parietal neurons area involved in hand preshaping simultaneously encodes multiple potential grasp movements before one is chosen for action (Baumann, Fluet and Scherberger 2009).

This is how MRs manage different complex motor scenarios and deal with several motor possibilities with respect to a single object, as well as different objects in different contexts - something addressed, but left unspecified by Pacherie.

2.4.5 Dorsal Multimodal Spatial Processing and MRs

I need to point out here some important characteristics of dorsal perception, neglected even by those claiming that MRs are only dorsal phenomena. I suggest that dorsal representations can also be conscious – a possibility denied by Pacherie, Jacob and Jeannerod, for whom there are no conscious action-guiding representations (Jacob and Jeannerod 2003: 252–255; see Brogaard 2011: 1094) (§ 2.1.1).

First, dorsal lesions disrupt the conscious awareness of the quality of objects in the peripersonal space (Gallese 2007; see also Jacob and de Vignemont 2010). This means that ventral perception is insufficient to obtain conscious (spatial) perception without dorsal processing (Rizzolatti and Matelli 2003), which even if functionally insufficient for normal visuospatial awareness, is nonetheless functionally necessary. Thus, conscious spatial processing depends on both streams. We have sufficient evidence for this claim. For example, posterior parietal and dorsal perceptions are crucial in the visual control of eye and hand movements in depth (Fearraina et al. 2009) and different families of AIP neurons encode the 3-D structure of shapes defined by binocular disparity linked to stereopsis (Romero et al. 2013)[21]. This suggests that dorsal-parietal projections process a lot of cues that seemed to pertain exclusively to ventral perception. This is an important implication for the relation between "higher" vision and dorsal vision. Also, many neurons in the rostral part of V-D are multimodal, sensitive to somatosensation, motor activity and visual stimuli in peripersonal space (Kravitz et al. 2011) – cfr. with (§ 2.3.3)[22] and crossmodal influences on dorsal processing suggest that perception-for-action is effectively multi-modal (Gentilucci et al. 1995) insofar as several posterior parietal spatial representations exhibit a multimodal nature (Andersen et al. 1997; Holmes and Spence 2004) – all this is in line with the idea that dorsal stream processing is definitely multimodal, as AIP is (Chinellato and del Pobil 2015). This is possible because V-D receives inputs from both streams providing - at least in the right hemisphere - the basis for higher-level spatial representation (Milner and Goodale 2006). This is because V-D integrates a variety of non-visual stimuli (including tactile, kinesthetic, proprioceptive) (Gallese 2007: 7; Kravitz et al 2011: 223, Box 1; Jeannerod

---

[21] Also, caudal intraparietal (CIP) neurons in the dorsal stream encode multiple depth cues (Tsutsui et al. 2005). For the crucial link concerning action, depth cues, stereopsis and egocentric localization see (Vishvanath 2014).

[22] Cfr. with footnote 19.

2006: 1.2.3; Fogassi and Luppino 2005) and transmits information to ventral perception required to construct conscious spatial representations (Bullier et al. 2001). Finally, V-D is crucial in maintaining attention to current task goals, and orienting to salient new stimuli (Singh-Curry and Husain 2009). This is an important specification for the peripersonal component of MRs, as well as its conscious counterpart, in the light of the open debate about the exclusive identification of the contents of visual awareness with ventral perception (Clark 2009; Schenk and McIntosh 2010).

Summing up, this second part of the chapter has clearly shown that MRs are subserved also by ventral vision (§ 2.4.3) – and are, in general, dependent on interstream interplay (§ 2.4.1), even though dorsal processing itself is more complex than widely believed in the philosophical debate. With my addition, we gain a lot of explanations concerning the functions reported in the decomposing strategy and we can add features to these functions that are usually neglected in the literature: MRs are not always automatic, but they can alternate between automatic and online processing and conscious or unconscious processing (§§ 2.4.3, 2.4.4, 2.4.5); they can rely on semantic computations (§ 2.4.3), due to their specific context dependency (§ 2.4.4). They can also be deceived, in some cases, by some kinds of illusion (§ 2.4.2), and can make use of both visual and motor memory (§ 2.4.1). But I want to suggest something very new for MRs. Also, in addition to what I said and against the view – which I criticized in the decomposing strategy - that MRs can be equated with dorsal perception without any specifications, it should be noted that, the dorsal stream can be divided not only into two sub-pathways, but, actually, into three sub-pathways – the divisions are compatible - all of them supporting conscious and non-conscious visuospatial processing, spatial working memory, visually guided action and navigation (Kravitz et al. 2011).

2.5 MRs: an Unexpected Functioning

Here I have a further important new point concerning MR processing. While dorsal perception distinguishes between images of graspable and non-graspable objects (Rice et al. 2007; Chao and Martin 2000), it cannot discriminate between normal and depicted objects, because this capacity is subserved by ventral perception: dorsal perception does not construct a *complete* 3D structural description of the target object. However, this description is necessary for response *selection*, in order to detect the action afforded by an object, or in the case of pictures, to understand that there is no possible interaction[23]. It is ventral perception that plays the key role in response selection, based on a comprehensive analysis of object volumetric structure, distinguishing between 3D objects and 2D images of objects by detecting conflicts between various visual cues[24] and selecting different visuomotor strategies for a 2D image versus a 3D object. Instead, dorsal perception plans the precise metrics of the intended action, based on a pragmatic analysis of the object's spatial features (see Westwood et al. 2002). For this reason, dorsal perception responds to depicted objects without the need of any volumetric representation, which is not possible in picture seeing[25]. Indeed, we have evidence that the visuomotor component (§ 2.3.1) is activated for depicted objects presented in the

---

[23] Volumetric object representation is necessary for the visual control of grip formation and response selection, to ensure that we do not attempt to reach for objects that cannot be grasped (Westwood et al. 2002).

[24] It is computationally efficient for one visual system to handle both response selection and object recognition: both require complete/detailed information about 3D object structure (Goodale and Milner 1992).

[25] For the difference between shape perception and volumetric object recognition see (Briscoe 2008).

peripersonal space of the observer (Zipoli Caiani 2013, Chao and Martin 2000; Romero et al. 2014; Costantini et al. 2010). This might be due to the fact that, in most cases, in the experimental settings the vehicle/surface of the depicted object (a monitor, a screen, a picture) actually falls within the peripersonal space of the observer as well and, since dorsal perception cannot really distinguish between a depicted object and a normal one once an object, whether depicted or real, is perceived - even if apparently - as located in the peripersonal space of the observer, dorsal perception responds. I'll be back to this in both chapter 4 and chapter 5.

It follows that, if MRs were not subserved by both streams, it would be hard for us to discriminate the nature of the objects we try to act upon, it being difficult to activate response selection and discriminate between normal and depicted objects. We are able to perform motor acts without clashing against a picture because MR is based on ventral/dorsal interactions. So this is another crucial argument for the dependence of MRs on the processing of both streams.

But there is a further crucial notion I want to point out. A common intuition is that, first ventral conscious perception selects the target/goal of interest for action and establishes if it is actable upon - if it is a real 3-D object, then, it sends the information to visuomotor dorsal perception which, only at this point, sets the parameters for interaction (see Young 2006: 140; Pacherie 2008: 186-187). That is, it is widely believed that visual consciousness establishes whether the computation of the coordinates for motor action can start due to the response selection establishing whether we are dealing with a reliable motor scenario (a real object) or not (a picture). However, things are exactly the other way. Indeed, the visuomotor encoding is activated, despite the fact - and even before - that ventral conscious perception has computed whether the object is "real" or not. Few indeterminate cues are in fact sufficient in order to trigger the visuomotor transformation: before ventral volumetric reconstruction a motor act computed on the basis of the 2-D geometrical properties of the target is already stored in our motor quiver, regardless of the fact that actual overt interaction will follow. So, our visuomotor system doesn't need any trigger, nor any confirmation from our conscious visual system in order to start the visuomotor translation. Rather, the translation is already ready to be used at the discretion of the ventral conscious response selection, which arrives later. This well explains the character of mental antecedents of action (Nanay 2013b, Jeannerod 2006) that MRs seem to have. The next section addresses an emotional component for MRs.

2.6 Extending the Neural Correlates of MRs even further: the OFC cortex

We saw that MRs extend beyond the two streams. Indeed, evidence shows that each stream projects to the orbito-frontal cortex (OFC), an area involved in emotional-affective encoding. This suggests an emotional component for MRs, crucially neglected in the literature[26].

During visual recognition, affective responses in the orbitofrontal cortex (OFC) – the orbital sector of the prefrontal cortex (Barrett and Bar 2009) - assist the visual perception of an object, integrating sensory input to build a contextually affective encoding of the world value (Barrett and Bar 2009). The two main OFC circuits connect with both visual streams: a medial OFC projecting to the dorsal stream computing initial

---

[26] To my knowledge, the only one that sketches the possible link is (Nanay 2013b: 155).

affective information and triggering the internal bodily changes suitable for action performance related to that object in context, by guiding autonomic, endocrine and behavioral responses. Due to connections with the lateral parietal cortex, the OFC's encoding of these autonomic and endocrine changes is relayed back to the dorsal stream as an initial affective estimate (Barrett and Bar 2009: 1329). Secondly, a lateral OFC projecting to the ventral stream, with robust reciprocal connections to the inferior temporal areas (TEO, TE and temporal pole) of the ventral stream (Barret and Bar 2009)[27] computing a broad gist affective computation of the object. Interestingly, the medial OFC encoding starts before the lateral OFC (for technical details, see Barrett and Bar 2009). This strengthens the presence of a privileged way from V1, through the dorsal stream, to the OFC and back to the dorsal stream. Therefore, since MRs are mainly dorsal phenomena, they are deeply linked to emotional encoding.

However, since both streams are connected to those emotional areas, the failure to address a further emotional component for MRs is due both to those who argue that MRs are dorsal phenomena and to those who argue that they are also the result of the interplay between the two streams as well. Thus, MRs emerge from a complex encoding, given by both streams and their projections to the OFC. Crucially, not only can MRs be conscious or unconscious, but also their related emotional encoding can, due to the link with the OFC (Barrett and Bar 2009). This is confirmed by behavioral evidence that while graspable neutral objects that can be approached without any risk activate a facilitating motor response, dangerous objects that pose a potential risk evoke aversive action possibilities, generating an interference-effect: information about an object's potential risks conflicts with the motor actions that are activated while observing that object (Anelli, Borghi and Nicoletti 2012). I cannot go into great detail here. Furthermore, this point is very important, insofar as affective perception might play a crucial role in shaping the process of competition of action possibilities I reported in (§ 2.4.3).

Accordingly, this emotional aspect of MRs is very important for a theory that aims to account for all the representational aspects of an MR. Here, I just sketched the basic empirical framework of this emotional aspect of MRs. Next chapter will precisely explain this important representational interplay between MRs and the representations of emotional relevant properties of the objects we deal with.

2.7 Conclusion of the Section

Summing up, I have reported sufficient evidence to defend the main twofold claim of the chapter: that MRs are not monolithic representational processes, but an ensemble of sub-representations with different functions and that those functions primarily rely on the dorsal stream but are also deeply dependent on interstream interaction, the ventral processing being crucial for the dorsal components of MRs. This specification is very important, insofar as, while in the neuroscientific literature it is widely agreed that, at a certain level of fine-graining, very few representational processes can be monolithic, the philosophical literature always talk about MRs without meticulously specifying their components and their complex anatomo-functional nature (see § 2.1.4). So, my theory aims to bridge the gap between philosophy and neuroscience concerning the notion of MRs and to establish their functions with respect to their neural underpinnings. I do that by reconciling our best philosophical theory about MRs with the most important sets of empirical evidence we get from vision and motor neuroscience, especially these about

---

[27] Remember that AIP connects with inferotemporal areas (TEm, TE and TEO).

motor perception and the TVSM, which are hotly debated in both philosophy and neuroscience. If we endorse my decomposing strategy, we are able to avoid the problems reported in (§ 2.1.4) and to add important features for MRs, which are usually neglected in the philosophical literature. This is also useful in order to build an account of MRs which can collect all the crucial empirical results coming from neuroscience, which, otherwise, risk being left out of our philosophical theory of MRs, remaining just a bunch of detached sets of evidence, with no background philosophical theory.

3 Motor Representations and Emotions

What is the role of emotions in action performance? Philosophy of action has never explicitly answered this question. We can note this by taking a look at some of the most important works on action (Bratman 1984, 1999, 2007, 2014; Bradie and Brand 1980; Brand 1984; Mele 1995, 2009, 2014; Anscombe 2000). Some authors just mention emotions very marginally (Bach 1978; Vargas and Yaffe 2014: 117), or directly admit to not taking them into account in relation to action (Mele 1992: 117-111). Others specify that we should talk about them (Hursthouse 1991; Mele 1995: 102), but no specific relation between emotions and action is proposed (Kenny 2003: Ch. 7). Emotions and action are coupled only when mentioning akrasia - a state of the subject in which s/he acts against his/her better judgement through weakness of will - and self-control (Mele 1995: Ch. 6, see also Mele 1987, 2003, 2009, 2014), or when talking about judgments (Mele 1995: 108-109-65-82), beliefs (p. 8), inferential errors (p. 86; see also Nisbett and Ross 1980: 228) and other mental states (Mele 1995: 106-122-138).

This omission arouses curiosity insofar as it seems natural to argue that emotions play some sort of role in our active evaluation of the environment. Indeed, the way we emotionally perceive the environment influences the way in which we manage our interaction with it. For this reason, establishing the relation between the mental antecedents of action and emotion is very important in order to explain the nature of the different motor situations we deal with. This is the main aim of this chapter.

First of all, note that emotions are largely absent in action theory (Zhu and Thagard 2002) because action is usually explained by *pro attitudes* like desires, intentions, or purposes, and epistemic factors like beliefs (Davidson 1963/1980) and the classic belief-desire models lack any explanation about the contribution of emotions to actions (Pacherie 2002).

An important point: not only does the belief-desire model lack any explanation about the relation between action and emotions. Also, this model, according to which beliefs and desires mediate between sensory input and motor output, has been recently judged incomplete for action explanation itself, insofar as, as we saw in the previous chapter, for most of the actions we perform, the only mental representational states mediating between sensory input and motor output are perceptual states named MRs, and not a set of beliefs and desires. Furthermore, even when beliefs and desires play an essential part in our actions, MRs still need to be involved (Nanay 2013b: 5, see 1.2; 2014a) in order to shape the content of an intention (Butterfill and Sinigaglia 2012; Pacherie 2000). The advantage of explaining action by referring to MRs is that they code simultaneously for things that are coded separately in the belief/desire model. In the belief/desire model, the psychological antecedent of an action includes both conative and cognitive elements: "on the conative side, a desire about a certain result; on the cognitive side, an orienting belief that one is in situation S and an instrumental belief that in S action A brings about R. Situation, goal and means are thus represented separately" (Pacherie 2002: 69). No such dissociation occurs with MRs (Ibid.). Thus, MRs shape the essence of action performance, being the representational component of the immediate mental antecedent of actions we need to have in order to perform a goal-directed action (Nanay 2014a: 4; Nanay 2013 b: 39, 2014a; Jacob and Jeannerod 2003, Jeannerod 2006; Pacherie 2000, 2002; Butterfill and Sinigaglia 2012).

All I have said suggests that we can explain the relationship between action and emotions by explaining the relation between emotions and MRs. This is also an advantage for a theory of MRs (Nanay 2013b; Butterfill and Sinigaglia 2012; Jacob and Jeannerod 2003) insofar as the literature on MRs is completely silent about an emotional component influencing them. I also said that, since MRs are not normally accessible to introspection, we need to turn to neuroscience in order to characterize and analyze them (Nanay 2013b, 2014a). This is a further advantage when it turns out that the influence of emotions on MRs may be not consciously accessible either and, for the same reason, only neuroscience can tell us something about this relation[28, 29].

In the previous chapter I just sketched a possible way of coupling between emotions and MRs. Here I push the line further. This chapter aims to explain the relation between MRs and emotions, by suggesting that the representation of object emotional-relevance properties (being dangerous, etc.), or emotional representations, can influence, through inhibition (or elicitation), the way in which we represent the action properties/possibilities (being graspable, climbable, etc.) inherent in the very same object: MRs can be emotionally charged.

Consider the following scenario: I am looking at a glass of water; it offers me an action property, the property of being grasped and an action possibility, the possibility of being grasped with a power grip. However, the glass is broken, thus my visual system attributes, through its projection to other cerebral areas subserving different mental representations, also another property to the glass: the property of being dangerous. Does the representation of the latter property influence the representation of the former? After all, even if the broken glass is graspable, I might cut my hand in trying to grasp it. If the glass is graspable, but the representation of the property of being dangerous is influencing my motor appraisal, I may decide not to grasp the glass. Hence, a sort of influence is in play here. Explaining this influence means to explain how MRs – and thus action - couple with emotions.

Before developing my idea, I have to mention two similar attempts. The first one is by Zhu and Thagard (2002), who suggested incorporating emotions within action theory for both the phases action can be divided into: the generation of an action (or action tendency) and the execution and control of an action. Unfortunately, their attempt is based on very old empirical evidence (LeDoux 1996: 164). Given the exponential growth of data in neuroscience, my account has the advantage of using recent results able to explain very carefully the role of emotion in the generation and selection of an action, which is the only action phase this chapter is interested in. I also sort out some problems raised by Zhu ans Thagard and Ledoux. The second attempt is by Pacherie (2002), who talks about MRs, but without giving an idea of how MRs couple with emotions at the cerebral level. Indeed, Pacherie wants to show how emotion can play a role in the causal theory of action. I precisely propose this coupling between MRs and emotions, in the light of evidence from neuroscience.

Another thing the reader should consider is that, differently from action theory, the connection between emotion and action (Arnold 1960; Frijda 1986; Lazarus 1991)[30] is present, to a limited degree, in emotion theory, where emotions are taken to be

---

[28] Since MRs are the mental antecedents of action, it is not problematic to see a possible link with action tendencies, which concern the generation of an action.

[29] Here I am not concerned with the conative component of action.

[30] See (Colombetti 2007, 2014).

evaluative processes. Evaluation is usually twofold (Lazarus 1991): a primary appraisal specifies whether the situation is relevant for personal well-being; a secondary appraisal analyzes possible ways of coping with the situation, such as "What, if anything, can be done about the situation?" (quoted in Zhu and Thagard, p. 27)". Someone has suggested that emotions do not impel actions directly (Zhu and Thagard 2002: sec. 8); they instigate the search for appropriate actions (Prinz 2004)[31]. A point everybody seems to agree on is that emotions are linked to tendencies to act (Frijda 1986: 71; Roberts 2003: 168, 159, 163; Elster 2003; Solomon 2003, 2004; de Sousa 2014). However, not only do emotions aid action, but they can also undermine action (Greenspan 1988; 2000; LeDoux, 1996:177, in Zhu and Thagard 2002: 29). While this suggests the presence of action in emotion theory, instead my account wants to vindicate a special place for emotion in action theory, something we lack at the moment.

In defending the idea of a representational influence between emotions and MRs, I maintain the classical philosophical terminology in saying that we perceptually represent certain entities as having certain properties, or perceptually attribute certain properties to certain entities. Of course, not all properties that we represent objects as having are perceptually represented. I perceive a cup of coffe as black, as spatially located and as big. But I can also represent this object as having the property of being the same cup of coffee I used yesterday to drink water; this would be a non-perceptual attribution or non-perceptual representation (Nanay 2011b). Consider also that another clarification is needed. Consider the following question:

c) What kinds of properties does our visual system attributes to objects?

As we have said, concerning MRs it is commonly agreed that, thanks to MRs, our visual system attributes action properties - and thus action properties is one of the possible answers to (c). However, it is important to distinguish (c) from a very different question:

(e) what properties does our perceptual system respond to/covary with in perception?

(c) is about which properties are attributed by the perceptual system (or what it represents entities as having) and not about which properties are out there (presumably in a causal relation with our perceptual system), as in the case of (e). The properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b). This distinction will be fruitful both in talking about MRs, and in the case of emotions (§§ 3.1, 3.5).

The chapter proceeds as follows. Since action has to be explained through MRs, whose investigation has to turn to empirical science, I underline the neural mechanisms pertaining to the activity of the ventro-dorsal stream (or V-D), which is the main cortical circuit for the MRs involved in the representation of action properties (§ 3.1). Since I have to couple emotions with MRs, I explicit the notion of emotion I am committed to and the neural correlate of emotion I refer to, namely, the orbito-frontal cortex (OFC) (§ 3.2). Then, I show that OFC is deeply linked to the dorsal stream (and to the V-D), which is the main neural correlate of MRs (§ 3.3). I then report behavioral evidence in line with the neural evidence (§ 3.4) and finally, on the basis of these sets of evidence, I offer a philosophical account explaining how emotional representations can influence MRs (§ 3.5).

---

[31] They beg for action without demanding it. Emotions are not output states, nor action commands, but perceptions of the body's preparation for action (p. 244).

3.1 The Ventro-Dorsal basis of MRs

In the previous chapter I showed how action can be explained through MRs, whose investigation has to turn to empirical science. We saw that one important component of the MRs is the visuomotor component (§ 2.3.1), which is crucially involved in the representation of action properties. In this chapter I mainly focus on this component of MRs, as well as on its neural correlates, when coupling MRs with emotional representations – though, of course, all the others components remain crucial. However, since the discussion will be framed in terms of the anatomo-functional underpinnings of the visuomotor component of MRs, it is worth recalling them here in order to fully understand the relation between them and neural underpinnings of the emotional representations I am going to talk about in this chapter.

As said, objects exhibit geometrical properties (e.g. size, shape, texture) that are, from a motor point of view, action/motor properties, to the extent that they permit a precise action possibility satisfiable with a precise motor act. For example the geometrical features of a mug are action properties permitting an action possibility (grasping), which can be satisfied by a proper motor act: a power grip. This important function of MRs is due to the visuomotor transformation that relies on a well defined parietal-premotor network lying in between the parietal cortex and the ventral premotor cortex, that is, a precise portion of the V-D, whose main components for these tasks are the anterior intraparietal (AIP) area and area F5, in the most rostral part of the ventral premotor cortex (for a review of the leading role of AIP-F5 in the detection of action possibilities[1] and the related visuomotor transformation of object properties in action properties and then of action properties into motor acts see Borghi and Riggio 2015, Raos et al. 2006, Romero et al. 2014, Castiello 2005; Castiello and Begliomini 2008 Chinellato and del Pobil 2015; Kandel 2013: Chap. 19 concerning lesion studies see Andersen et al., 2014; Turella and Lignau 2014)[32]. AIP is one of the end-stage areas of the dorsal stream. AIP neurons respond selectively to objects during both passive fixation and grasping, extracting visual object information concerning action possibilities for grasping purposes (for a review see Romero et al. 2014, Raos et al. 2006); then, they relay this information to F5 neurons, with which AIP is directly connected (Borra et al. 2008), which then activate the primary motor cortex. In F5 we find visuomotor canonical neurons, which use the information received by AIP about action properties of the objects and compute the motor commands in order to interact with them. Also canonical neurons respond during object fixation, regardless of the actual execution of an action. In canonical neurons activity, there is a strict congruence between their high selectivity for a particular type of prehension (executed grip) and the visual selectivity for objects that, although differing in shape, require the prehension in order to be grasped (for a review see Raos et al. 2006).

Summarizing, I want to recall the example I reported above that when I am looking at the cup of coffee on my desk, V-D responds to those 3D geometrical properties of objects that serve such visuomotor tasks as grasping them. The AIP-F5 parietal-premotor network is crucial in translating those geometrical features into action properties and then into motor acts. Then we can grasp the cup. First of all, AIP detects the geometrical features of the handle that exhibit precise motor characteristics. This

---

[32] Empirical data – that I cannot review here - suggest the leading role of (both AIP-F5 and VIP-F4 in the) V-D concerning the visuomotor transformation (in the peripersonal space) (for a critical review see Turella and Lignau 2014; Kandel et al. 2013: 871). For coherence toward neuroscience, I specify that, obviously, my analysis is not so technical as those found in the neurophysiological literature (Borghi and Riggio 2015). However, what I say here is sufficient in order to defend my point.

means that shape, texture and size are encoded. Thus, the geometrical features are read as action properties. This information is sent to F5, which, given the information received by the AIP, computes the most suitable motor act (with respect to my motor repertoire, say, a power grip) in order to grasp the handle of the cup. At the same time, during this translation, the simulation of the appropriate motor behavior is encoded: the visuomotor brain is both perceiving the action possibility and simulating its readiness for potential related motor interaction, setting a motor act "in the quiver"[33].

Now, coming back to what I have said before, it is important to distinguish
(c) What kinds of properties does our visual system attribute to the object?

from a very different question

(e) what properties does our perceptual system respond to/covary with during perception?

(c) is about what properties are attributed by the perceptual system (or what it represents entities as having) and not about what properties are out there (presumably in a causal relation with our perceptual system), as in the case of (e). The properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b). I also said that the answer to (e) is fundamental to get the answer to (c). I'll be back to this distinction in chapter 4.

The neural phenomena described above translate object attributes into motor commands: precisely, they represent S (object geometrical features) as being F (action properties) and F (action properties) as Q (motor acts) resonating with the simulation of the performable motor act. But I have said that the properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b). Concerning the visuomotor transformation process, our visual system responds to/covaries with particular "object geometrical features". Thus, the information goes to the V-D, which reads those properties as action properties. In particular, I said that the properties our visual system responds to or tracks may not be the same as the ones it represents objects as having, those it attributes to objects (Nanay 2011b). Thus, the properties our visual system responds to or tracks (geometrical properties) are not the same as the ones it represents objects as having (action properties), through the V-D processing. Yet, the detection of particular geometrical features allows these geometrical features to be transformed into (or read as) action properties, or, in other words, it allows the attribution of action properties to the object. Accordingly, while object geometrical properties are out there, action properties are core-relational properties. This is accompanied by the motor resonance of both the two main components in the covariation and attribution: AIP-F5. Thanks to simulation, it is as if the brain were encoding that a performable motor act is "in the quiver".

As I said in the previous chapter, it's worth pointing out that the visuomotor transformation performed by MRs also needs the representation of a goal (Butterfill and Sinigaglia 2012, Pacherie 2000, Jacob and Jeannerod 2003) (my § 2.3.2) and MR correlates also involve a semantic contribution from the ventral stream (Rizzolatti and Sinigaglia 2008: 36-38; Fogassi and Luppino 2005: 627; Young 2006: 134; Borghi and Riggio 2015) (my §§ 2.4.1, 2.4.3), so that MRs can be considered as subserved by both streams (Nanay 2013b: 3.4) due to important evidence of interstream interaction at every stage of the visual processing (Schenk and McIntosh 2010), as meticulously explained

---

[33] For the behavioral counterpart of these experiments see (Borghi and Riggio 2015; Chinellato and del Pobil 2015).

before in my (§ 2.4)[34]. So, on the one hand, I take the visuomotor transformation, and its dorsal neural correlates, the most important feature of MRs, insofar as it explains the mechanisms through which the visuomotor brain transforms the visual input in motor output, by building several MRs for potential motor interactions and this is a rigorous model for the explanation of the mental antecedents of action (Fadiga et al. 2000; Cisek 2007; Cisek and Kalaska 2010; Kandel et al. 2013: Chap. 19); on the other hand, I maintain that, while MRs are mainly dorsal phenomena, they can be subserved by both streams, as well as other cortical projections (see Nanay 2013a, b for a review)[35] – that is, I maintain what I said in developing my theory of MRs in the previous chapter.

Since I have just introduced the notion of MRs and I intend to couple them with emotions, I have to say something about what I mean by "emotion".

3.2 Emotions

Philosophically speaking, I assume, as most of the scholars of the literature do, that emotional states are representational states, they refer to something, they have content (Peacocke 2001; Barlassina and Newen 2013; Wollheim 1999; Charland 2002; Prinz 2004; Tappolet 2003; de Sousa 2014; Goldie 2000; Nanay 2013b), that they attribute salient emotional relevant properties to objects (I owe this expression to Nanay 2013b: 155; but see also Barlassina and Newen 2013) and, following those who attempt to naturalize emotions, that they depend on neurophysiological states, which is not controversial (Barlassina and Newen 2013; Prinz 2004; Roberts 2003; Doring 2003; de Sousa 2014; Nanay 2013b). The usual loci where we can detect emotional activity in the brain are the insular pathway (Craig 2002, 2003), the somatosensory pathway (Khalsa et al. 2009) and the orbito-frontal sector of the prefrontal cortex, that is, the orbito-frontal cortex (OFC), connected with the amygdala and the ventral striatum (Barrett and Bar 2009). Here I mainly focus on the OFC, which is crucial for emotional encoding (Duncan & Barrett 2007; Pessoa 2008), but the reader should know that these different cortical states are deeply interconnected through several projections (Barrett and Bar 2009: 1326).

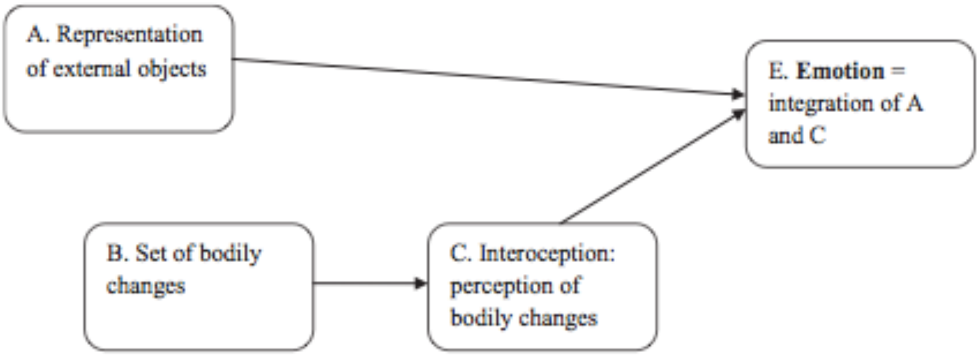
An important point with the naturalization of emotions is to establish their content. A common view about emotions is that they are interoceptive states (bodily perceptions) of bodily changes (Prinz 2004: 5, 58). But, if emotions are interoceptive states, then, one has to explain how they are directed towards particular objects. The usual strategy is to talk about co-occurrence: "the neural representation of an emotional bodily state fires at the same time as the neural realization of the representation of its particular object" (Prinz 2004: 181). This notion presents a problem because emotions are, then, directed towards a myriad of things. But this should not be the case: though my fear of the tiger co-occurred with many representations, it was directed only towards the tiger. In order to avoid this problem, Barlassina and Newen (2013) suggested that emotions (E), are twofold states given by: (C) interoceptive states and (A) representations

---

[34] AIP is interconnected with inferotemporal areas (TEm, TE and TEO) of the ventral stream. Tem neurons process object recognition through 3D visual cues and AIP is involved in both inferior temporal and posterior parietal activity for the three-dimensional shape perception (Fogassi and Luppino 2005).

[35] This endorsement is important because, in what follows, I show the connection between MRs neural correlates and the representations performed by the emotional area OFC. We will note that, while the dorsal stream, and in particular, the ventro-dorsal stream, are deeply connected to the OFC, several projections of the ventral stream to the OFC pass through the connections between the ventral and the ventro-dorsal stream. This is the empirical ground for the claim that emotional representations can influence MRs and, thus, that MRs can be emotionally charged.

of external objects (events, or states of affairs)[36]. (A) provides an emotion with its particular object. The emotion of fear expressed by the perception of bodily changes is directed towards/represents that snake because it has the perceptual representation of that snake among its constituents as being dangerous, or has having a somewhat emotional relevance property[37] - below the schema provided by the authors (p. 26).



Endorsing this notion of emotion, I now want to provide empirical evidence concerning the anatomo-functional link between the neural correlates of MRs reported in (§ 3.1) and the OFC, which is the example of emotional area I will focus on later. This will be the empirical ground in order to assess a philosophical explanation of the relation between MRs and emotions in the last part of the chapter.

3.3 Visual Perception for Action and Emotional Charge: Neural Evidence

I said that MRs are mainly dorsal phenomena, while I endorse the presence of a ventral component for their processing (see chapter 2). Here I will show that these neural correlates are deeply interrelated with a cortical area involved in emotional encoding, namely, the OFC. This will be the crucial empirical background in order to establish the relation between MRs and emotions by explaining the link between their neural correlates.

First of all, we know that visual perception is always emotionally charged at an early visual stage because affective computations signaling an object's affective impact and relevance assist the visual perception of an object from the very beginning of visual stimulation. Thus, affective and visual aspects of the environment are encoded simultaneously. This affective encoding is mainly due to the OFC (Barrett and Bar 2009), which activates between 80 and 130 ms after stimulus onset during the presentation of isolated objects and integrates sensory input from the world and the body to build a contextually sensitive, multimodal encoding of the world and its value.

However, we can distinguish two functionally related OFC circuits differentially connected to the dorsal stream and to the ventral stream, with different roles for affective encoding during object perception (I meticulously follow here the model proposed by Barrett and Bar 2009).

---

[36] Given this, one might argue that using the term emotional representation could be redundant. But since it is useful to recall the representational nature of emotions, I will use the expression emotional representations when talking about the representational counterpart of emotions.

[37] I will use the expressions emotion and emotional representations interchangeably.

The Medial OFC projects to the dorsal stream and has strong reciprocal connections to its lateral parietal areas (MT and MST). Through largely magnocellular pathways, the medial OFC receives the same low spatial frequency visual information, devoid of specific visual detail, that is used to create a basic-level category encoding of the object's identity (Barrett and Bar 2009) and computes initial affective information about what an object means for a person's well-being, triggering the internal (autonomic, endocrine) bodily changes suitable for action performance related to that object in context (Price 2007). Due to its neuroanatomical connections to the lateral parietal cortex, the OFC relays back this information about autonomic and endocrine changes with the information about the spatial location of the object to the dorsal stream as an initial estimate of the affective value and motivational relevance (Barrett and Bar 2009: 1329). The brain's preparation to respond (based on this gist-level prediction) arrives even before the object is consciously perceived. This shows the intrinsic influence the emotional response might have on the motor response.

What is important to point out in the light of these sets of evidence is that, since MRs are mainly dorsal phenomena, the link between these emotional areas and the dorsal stream is an important point in the way of establishing an emotional shade for MRs processing. Indeed, the medial OFC processing, with the related bodily preparation, starts before the lateral OFC. Following Barrett and Bar (2009: 1330), "visual information arrives more quickly to the medial OFC owing to a 'magnocellular advantage' in visual processing (Laylock et al. 2007). Magnocellular neurons projecting from the lateral geniculate nucleus (in the thalamus) rapidly conduct low spatial frequency visual information to V1 and the dorsal stream areas, compared with the parvocellular neurons carrying high visual spatial information to V1 and to the ventral stream. In humans, magnocellular neurons in V1 fire from 25 ms (Klistorner et al. 1997) to 40 ms (Paulus et al. 1999) earlier than parvocellular neurons in V1. Also, sometimes some neurons within the dorsal stream that receive input directly from the lateral geniculate nucleus (e.g. V5[38]/MT; Sincich et al. 2004) become active even before V1 (Ffytche et al. 1995; Buchner et al. 1997). Also, neurons in the prefrontal cortex become active at approximately 10 ms after neurons in the dorsal stream, but are coincident with the activation in the ventral stream (Foxe and Simpson 2002)". This evidence seems to show that the affective computation is so fast that we cannot detach this from the processing of MRs. This also shows the presence of a privileged way that goes from V1, through the dorsal stream, to the OFC and back to the dorsal stream: vision-for action entertains a privileged emotional nature.

But, we also know that the MRs visuomotor transformation subserved by the ventro-dorsal stream relies also on ventral perception, so that MRs can be considered a product of the processing of both streams. To this extent, note that there are parvocellular pathways connecting the lateral OFC to the inferior temporal areas (TEO, TE and temporal pole) of the ventral stream (Barret Bar 2009 for a review) with which the ventro-dorsal stream is reciprocally interconnected (Fogassi and Luppino 2005) – see foonote 35.

To sum up, the OFC processing entertains reciprocal influences with both the ventro-dorsal and the ventral stream. But, since MRs are subserved by both streams (Nanay 2013b), this suggests that, concerning the neural correlates, MRs are linked to emotional encoding. The next section reports related behavioral evidence. Both these sets of evidence will be crucial, in (§ 3.5), to explain the relation between MRs and emotions.

---

[38] It is important to know that V-D (related to the inferior intraparietal lobule, IPL, Gallese 2007) — of which AIP is part — receives direct inputs from the visual areas MT/V5, which in turn receive inputs from the primary visual cortex V1.

3.4 Behavioral Evidence

In line with (§ 3.3), we have behavioral evidence that observing objects elicits facilitation effects of motor responses about action preparation (Anelli, Borghi and Nicoletti 2012) with respect to the object we are confronting. Indeed, while neutral graspable objects that can be approached without any risk activate a facilitating motor response, with dangerous objects that pose a potential risk, motor resonance evokes aversive motor responses, generating an interference-effect. That is, the information about an object's potential risks might conflict with the motor actions that are activated while observing that object and then it is blocked, when aversive action possibilities are evoked. Indeed, due to this interference-effect, response times are slower with dangerous objects than with neutral objects.

This evidence is compatible with data revealing that response times are generally slowed down with emotional stimuli due to an inhibition effect provoked by a selective attention mechanism (Algom et al. 2004) and that the prefrontal cortex plays a double role, exerting both an inhibitory and an excitatory control[39] (Munakata et al., 2011), allowing participants to inhibit the tendency to respond to action possibilities in the case of dangerous objects (Caligiore et al. 2013). Also, somatosensory cortices/IPL, anticipate the consequences of observed hand-object interaction with noxious objects, as the painful grasp condition is activated to a greater extent compared to all other conditions (Anelli, Borghi and Nicoletti 2012: 1637).

The same evidence comes from the study of broken or dangerous action possibilities. For example, the action possibilities offered by a broken handle are not activated, or are activated and then inhibited, due to a weaker activation of the cortical areas typically involved in performing action when the handle is intact (Buccino et al. 2009; Riggio et al. 2006; Anelli et al. 2013a, b; Ohman et al. 2001; Algom et al. 2004; Caligiore et al. 2010, 2013).

This evidence clearly shows that the affective valence of objects influences motor behavior and is in line with evidence reported in (§ 3.3). Following the evidence in (§§ 3.3, 3.4) I will now develop my philosophical proposal about the representational interference between MRs and emotions.

3.5 MRs and Emotions

Here I develop my idea of a representational intereference between emotion and MRs. First of all, I have to clarify the relation between OFC activity and the idea of emotion exposed in (§ 3.2). Then, I can develop my explanation (§ 3.5).

In (§ 3.2) I embraced a widely agreed idea of emotion and listed the neural correlates of emotions mentioned in the literature. The emotions I talk about here are those whose neural correlate is the OFC. That means, following the idea of emotion proposed by Barlassina and Newen (§ 3.2) - which is able to avoid several problems emotions encounter in the emotion theory - that OFC is a neural state at the basis of an interoceptive state dealing with the perception of precise bodily changes plus the representation of an object. This representation, linked to interoceptive states, represents the emotional relevant/salient properties of the object.

---

[39] During simulation, muscular contractions are effectively present, but simultaneously blocked by inhibitory mechanisms (see Jeannerod 2006: 2.3.3).

I distinguished questions (c) and (e) (§ 3.1) when talking about what MRs represent. But the activity of the OFC emotional states described in (§ 3.3) functions in a similar manner with respect to MRs, in relation to these questions. On the basis of the covariation of my visual system with respect to the object geometrical properties, the visual information goes to the OFC, which reads these properties as emotional salient properties. But the properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b). Concerning the emotional representations, our visual system responds to/covaries with particular "object geometrical features". Accordingly, on the basis of this covariation, and of the information arriving from the visual system, our OFC attributes "emotional-relevance properties" to the same object. However, while object geometrical properties are out there, emotional-relevance properties are core-relational properties (for the idea of core-relation see Prinz 2004).

So, I specified a similar answer, concerning both emotions and MRs, concerning (c) and (e). Starting from this representational isomorphism between OFC states and MRs in relation to questions (c) and (e), I can explain how MRs are influenced by emotional representations, namely, by the representation of emotional relevant/salient properties.

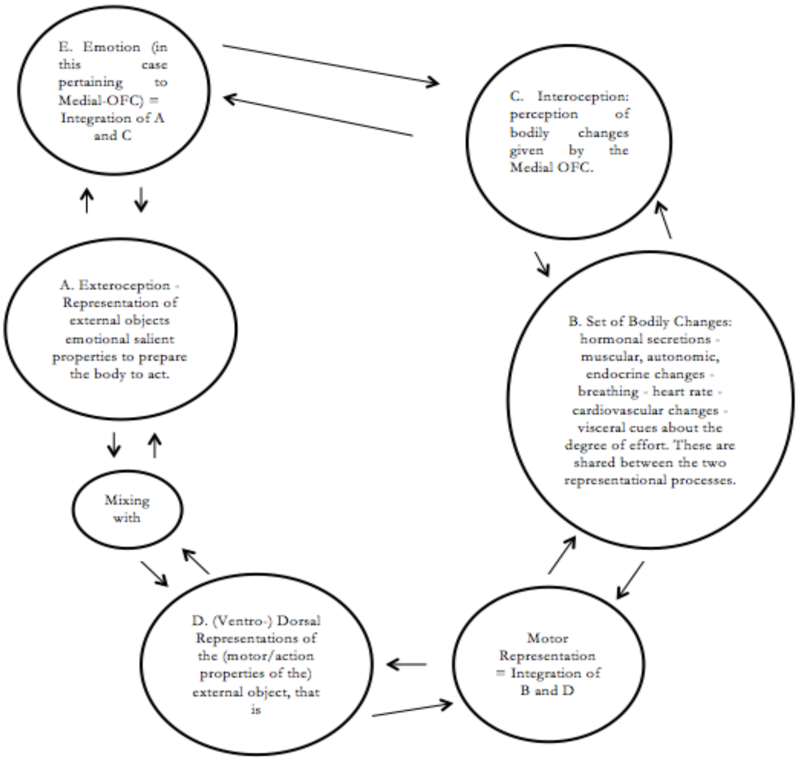
Indeed, I said that emotional representational states represent emotional relevant properties, while MRs represent action properties and the related motor acts. Saying that an MR can operate in an emotionally charged manner means that the ascription of relevant-emotional properties from an emotional representational state can influence the ascription of action properties due to the MRs. For example, I can represent a graspable object as dangerous; that is, I can represent an object as both graspable and dangerous; or, I can represent that object as dangerous and thus non-graspable.

Twofold representational ascription is possible because there are some situations in which, though our visual system is covarying in its whole with the geometrical features of an object, our brain ascribes, on the basis of this covariation, two different kinds of properties to the same object, thanks to the information flowing between two different cortical states our visual system projects to: in this case, the V-D activity subserving the MR ascribes an action property (which is possible to ascribe also to a broken object), and the OFC activity subserving the emotional state ascribes an emotional-salient property. This is not controversial: we usually entertain parallel but different cortical/mental representations of the same object, while these two different representations may conflict. For example, we can entertain a dorsal visual and a ventral visual representation representing the same chip as having different size properties in the case of the 3D Ebbinghaus illusion (Nanay 2014b). Given this, while in the previous sections I provided an answer to (c) and (e), I will now explain, in the light of these answers, the overlap between the processing of MRs (§ 3.1) and emotional representations (§ 3.2).

First, concerning MRs, the visual system covaries with geometrical properties of the object, and, thanks to the information reaching the V-D, these geometrical properties are read, by the MR, as action properties, which trigger the relative motor acts: on seeing the object we obtain at the same time its visuomotor priming (i.e. the visuomotor representation of its "action property"), and the internal *simulation* of one of the actions we could perform upon it (i.e. the most suitable motor program required to interact with it with respect to the action possibility), regardless of actual action execution (§ 3.1). But crucially, at the same time, on the basis of the covariation of the visual system with respect to the same geometrical properties of the object, the information goes to the OFC, which reads those properties as emotional salient properties, pertaining to the same object; say, a broken handle of the very same mug (§ 3.4).

This means that the same geometrical property in the environment can be read by two representational states, namely an MR and an emotional representation, as instantiating two different properties: an action property and an emotional-relevant property. What is important to stress, for the purpose of this chapter, is that those two "representational readings" march in step, influencing each other reciprocally. For example, the represented action property in the case of a damaged object can lose its valence when my emotional state ascribes the emotional salient property of being dangerous[40] which allows me to perceive the action property in an emotionally charged manner, (e.g. I am afraid of cutting my hand with that dangerous object if I satisfy the action property).

But we should not forget that emotional representations and MRs march in step with bodily changes linked to motor simulations (see § 3.1) and emotional responses (see § 3.3), which share similar bodily activities: changes in heart, respiration, and cardiovascular rates during motor simulation, along with the representation with visceral cues about the degree of effort involved in the represented action (Jeannerod 2006, Ch. 2), are quite compatible with the bodily changes due to OFC. Most of the bodily changes guiding motor inhibition and emotional inhibition are very similar. Then, it is reasonable to suppose that the OFC representation of a salient property (of being dangerous), which triggers a particular bodily state, inhibits the bodily changes motor simulation is accompanied by. Indeed, as I have pointed out throughout the chapter, both OFC and MRs states not only represent something in the environment, but are also linked to precise bodily states. The description of my model is shown in the following schema.



---

[40] … which, in a certain sense, can be seen as another action possibility, the possibility for someone to cut. Thus, the representational influences of a representation sometimes lead to switching the nature of the attributed property of the other representation, even if the other representational device in question deals with particular properties (e.g. action properties).

Here's an example. Imagine you have to grasp a cylinder whose surface is coated by spurs. The slots between the spurs are more or less as large as your fingers. The spurs are very sharp. The first time you are faced with the cylinder, the perception of the action possibilities given by the fissure will be accompanied by perception of this object as dangerous; we perceive an action possibility for this object in a "scared" manner. However, if we carefully practise laying our the fingers between the slots, so that, when we see the object, our brain can immediately compute the grip aperture[41] in order to make the hand fit with the dispositions of the spurs, the object will be perceived in a different emotional manner, namely, in a less scary manner with respect to the first time. Thus, the represented motor act won't be accompanied by the ascription of the same range of danger with respect to the initial attempts. The accuracy of our MRs has changed, but what has also changed is also the influence on them by the representation of the emotional relevant properties we ascribe to the object: they march in step. In behavioral terms, the priming of our motor response will not trigger a deep inhibition as in the first attempts. In neural terms, once our visual system has detected the geometrical properties of the object, two neural states will construct different representations influencing each other and, this time, the gist level prediction from the OFC to V-D, will not be as negative as at the beginning. This further strengthens the link between represented action properties and represented emotional-relevant properties.

3.6 Conclusion of the Section

I think that what I have said is sufficient for arguing that the representation of object emotional relevant properties can influence, through inhibition (or elicitation), the representation of an action property of the same object.

The way we emotionally perceive the environment influences the way in which we manage the interaction with it, insofar as emotions influence our MRs, which are the constitutive mental antecedents of action. If we avoid this important emotional feature of these mental antecedents, we are not able to explain a wide range of situations we encounter when we try to investigate action.

In the next chapter I defend the idea, completely new, that the representation of action properties, which is a function of MRs, in particular of the visuomotor component (§ 2.3.1), is possible not only with normal objects, but also with depicted objects.

---

[41] An automatic process of grip formation taking place during the transportation of the hand in which the fingers are preshaped much before the hand touches the object. At about 60% of its transportation the hand reaches its widest opening, or *maximum grip aperture* (henceforth: MGA) and the size of the finger-grip at MGA (though much larger than the object to be grasped) is linearly correlated with the size of the object (Jeannerod 2006: 5).

4 Pictures, Action Properties and Motor Related Effects

The most important question concerning picture perception is

(1) What perceptual state are we in when we see an object in a picture?

Here I maintain the classical philosophical terminology used in the literature in saying that our visual system uses perceptual representations and thus perceptually represents certain entities as having certain properties; it perceptually attributes certain properties to certain entities[42]. So, (1) can be reformulated:

(1) what kinds of properties does our visual system attribute to the depicted object?[43]

In order to answer (1), philosophers have used the results of the TVSM, according to which our visual system can be divided into two streams, a ventral stream for object recognition, allowing one to perceive from an allocentric frame of reference, and a dorsal stream for visually guided motor interaction, thus allowing one to perceive from an egocentric frame of reference. The dorsal/ventral account of picture perception (see Nanay 2015) was first proposed by Matthen (2005) and developed by Nanay (2011a). Following this model, the authors denied that we can be in a dorsal perceptual state when perceiving a depicted object. This is because a depicted object is not physically graspable or manipulable and, in turn, it cannot be egocentrically localized, as a normal object, by the dorsal stream. Thus, the impossibility of manipulating depicted objects and of localizing them from an egocentric frame of reference has led some people to be sceptical about the possibility of a representation of action properties in the perception of objects in pictures, which pertains to the dorsal visual system.

Summarizing, those who embrace the empirical results of the TVSM in order to answer to (1) seem to conclude that since our visual system cannot egocentrically represent the depicted object, we cannot represent its action properties, which is the main task of our dorsal stream. So, one of the big differences – and the most important one in the dorsal/ventral framework - between face-to-face perception and the perception of the depicted objects would be that we (our dorsal stream) can attribute action properties to real objects, but not to depicted objects[44].

The aim of the present chapter is to show that, among the possible answers to (1)[45], there is the precise claim that it is possible for the depicted object to be represented by dorsal perception. That means that we can ascribe action properties to depicted objects as well. The dorsal ascription of action properties depends on what I call the visuomotor component of MRs, and its neural correlates (§ 2.3.1) and on its relationships

---

[42] I am not defending a representationalist point of view at the expense of the anti-representationalist one. Everything I say here can be reformulated in anti-representationalist terms (e.g. presentation).

[43] Of course, not all properties that we represent objects as having are perceptually represented. I perceive a cup of coffee as black, as spatially located and as big. But I can also represent it as having the property of being the same cup of coffee I used yesterday to drink water, and this would be an example of how it is possible to attribute properties non-perceptually to objects, or, to represent them non perceptually (Nanay 2011b).

[44] Of course, in answering (1) it does not seem strange to argue there are properties that can be perceptually represented in both real object perception and picture perception (colors, shapes, etc.) while other properties can be attributed only to real objects – or picture surfaces, which are real objects (the spatial localization in egocentric coordinates) – but not to depicted objects.

[45] I mean among those that can be given in the framework of the dorsal/ventral accounts of picture perception.

with the egocentric/peripersonal component and its neural correlates (§ 2.3.3) – I leave out the goal related component here. So, I will recall them along this chapter, but only in order to explain how they can be involved in picture perception as well.

At first sight one might think that my attempt is implausible, because it is widely agreed, also by commonsense, that a depicted object cannot be physically grasped. So why argue that the main cortical circuit involved in the motor interaction with objects is also involved in the perception of depicted objects, which clearly cannot be grasped or manipulated? In order to properly understand what I am arguing for and what this chapter is about, think about the following situation: I am looking at the cup of coffee on my desk. I represent it as having action properties, given by its geometrical features and, in particular, its handle offers me an action possibility: for example, grasping it in a power grip, thanks to its geometrical configuration. However, there's no more coffee in the cup, so I decide to get another one from the coffee machine. I look at the coffee machine and I notice on its cover a picture of a cup of coffee of the same kind as the machine delivered to me half an hour before. So I wonder: what is the relation between the perception of the real cup of coffee on my desk and the perception of the picture of the cup of coffee on the cover of the coffee machine with respect to the representation of action properties? Can the latter physically afford action possibilities pertaining to real interaction like the former? Of course it cannot, and, of course, I am not committed to this idea. Indeed, it seems trivially true that, from a physical point of view, we cannot grasp a depicted object. But does the physical impossibility of manipulating depicted objects mean that my visual system cannot attribute action properties to them? Following the dorsal/ventral account of picture perception, the answer seems to be positive.

Here I resist this conclusion by arguing that the same dorsal perceptual representation of action properties is shared by both face-to-face perception and the perception of depicted objects. Indeed, the fact that we cannot physically grasp a depicted object does not imply that our visuomotor brain cannot resonate to the action possibilities (simulating, for example, a power grip) fostered by the representation of action properties linked to the geometrical features of an object when the object in question is a depicted object. Hence the argument of the chapter. I report and discuss different sets of empirical evidence from motor and vision neuroscience showing that: first, dorsal perception can represent depicted objects as having action properties, insofar different components of the dorsal stream involved in the representation of action properties during the perception of normal objects are active during the perception of depicted objects as well; second, this is possible when, in the experimental settings, the depicted object is perceived as apparently located within the subject's peripersonal space and as apparently reachable (technical details below); third, this is strengthened by evidence that dorsal perception cannot really distinguish between a depicted object and a normal one. This means that, once an object, whether depicted or real, is perceived - even if apparently - as located in the peripersonal space of the observer, dorsal perception entertains a non-trivial egocentric relation with the target, and responds to the apparent motor possibility fostered by the distal target, regardless of the distinction real/depicted; a distinction indeed, that dorsal perception cannot draw.

Note that, philosophically speaking, my proposal has a crucial implication for the dorsal/ventral account of picture perception, as well as for the debate on pictures in general. If I am right, my account represents a sort of closure of the circle. Indeed, both streams can represent both the depicted objects (and their surfaces) and the normal objects in a very similar way, contrarily to what has been claimed. But if those two streams are the total components of the visual system of humans (and other mammals), then, our visual system functions in quite the same way in both picture perception and face-to-face perception. Thus, pictures and normal objects do not differ that much for our twofold visual elaboration, as the accredited model suggested, that is, concerning the representation of action properties. On the one hand, this doesn't mean that face-to-face perception and picture perception are precisely the same perceptual phenomenon; on the other hand, it means that they are more related than suggested by the dorsal/ventral account I am reviewing here. Moreover, my review represents an extension of this account, because I can address those cases that suggest this deep similarity, which have not been addressed before, while remaining in the same philosophical framework.

However, here I am not interested in the general literature on picture perception - which is trying to define the differences between picture perception and face-to-face perception. Again, I am not interested in the twofold experience of what Wollheim (1998) called seeing-in, namely, the relation between the perception of the picture surface and the perception of the depicted object (see Nanay 2011a). I am interested only in the latter term of this relation, while leaving aside the former. As I said, I am concerned with a precise portion of the literature, namely, with those accounts that use the dorsal/ventral distinction in order to answer to (1). But, again, my interest is for a particular claim concerning a particular issue raised by those accounts, namely, the issue concerning the (im)possibility of having a dorsal representation of action properties in the perception of the depicted object. Finally, note that this chapter deals with the case of normal pictures, leaving aside the case of particular pictures involving *trompe l'oeil* (see Nanay 2015 for the distinction dorsal/ventral with respect to those particular cases).

Before defending my claim, I need to introduce the dorsal/ventral account of picture perception.

4.1 The "Two Visual Systems Model" meets the debate on Picture Perception

The dorsal/ventral account is the best empirically framed philosophical account of picture perception in the light of vision neuroscience, which reconciles philosophy with neuroscience (Nanay 2015). This is because it applies the neurophysiological knowledge about vision we get from the TVSM in order to answer the main questions concerning the debate on pictures - especially question (1) - (see Nanay 2001a, 2015). Accordingly, despite the wide discussion about the implications of this model[46] in the case of normal objects, however, it is the best neuroscientific account we have about vision (Kandel et al. 2013). It was Matthen (2005) who first suggested the possibility that the dorsal/ventral distinction may be relevant for understanding picture perception (see Nanay 2011a: 477). Recall that, in a nutshell, according to the TVSM, there is a separation of the visual pathways, grounded on distinct anatomical structures (Milner, Goodale 1995): one for visual representation or experience, and one for visually guided action. It is well known that, accordingly, these paths can be dissociated due to cortical lesions. Lesions in the dorsal stream (from the primary visual cortex to the posterior parietal cortex) impair one's ability to use what one sees to guide action (optic ataxia), but not object recognition; lesions in the ventral stream (from the primary visual cortex to the infertemporal cortex) impair one's ability to recognize things in the visual world (visual agnosia), but not action guidance (see Jacob and Jeannerod 2003). That is to say, according to this model, the ventral stream is responsible for identification and recognition, allowing one to perceive from an allocentric frame of reference, whereas the

---

[46] There is an ongoing debate about the effective dissociation (the literature is too large to survey, see Nanay 2013a, 2015; Briscoe 2009). Here I do not use this argument to undermine in principle the dorsal/ventral account of picture perception.

dorsal stream is involved in guiding action, thus allowing one to perceive from an egocentric frame of reference.

A complete dorsal/ventral account of picture perception in this style, employable in answering (1), has been proposed by Nanay (2011a). He argued that: (a) since the perceptual representation of the depicted object is grounded on our recognitional apparatus, which in turn is grounded on the ventral subsystem, the depicted object is represented by ventral perception; (b) the depicted object is not represented by dorsal perception; (c) the picture surface is represented by dorsal perception; (d) the surface of the picture is not necessarily represented by ventral perception. In other words, for picture perception it is necessary that our ventral vision attributes properties to the depicted object, whereas our dorsal vision attributes properties to the picture surface. When we see an object face to face, our dorsal and ventral visual subsystems attribute properties to the same object: the perceived object. When we see objects in pictures, in contrast, the dorsal and the ventral visual subsystems attribute properties to different objects. The ventral subsystem attributes properties to the depicted scene, whereas the dorsal subsystem attributes properties to the surface of the picture (pp. 466-477).

While Matthen holds that (b) is necessary, Nanay simply holds that while he is not committed to the idea that (b) is of necessity true, he accepts that, most of the time, it is the case. Note, however, that the fact that for Nanay's (b) is not necessary becomes clearer only with his recent specification in (Nanay 2015)[47].

Indeed, in his (2015), Nanay specifies that he does not actually claim that it is *impossible* that the depicted object is represented dorsally[48]. It might be the case, he says, but normally it is not[49] (Nanay 2015). So, considering the question of whether our dorsal stream represents the depicted object for Nanay, "it is important to point out that at least in the case of normal (non-trompe l'oeil) pictures, the answer is *usually* (emphasis mine) negative: when we see Mona Lisa, the depicted face is unlikely to be represented dorsally. Dorsal vision is supposed to help us perform perceptually guided actions. But we don't and can't perform actions on depicted objects. Further, a minimal condition on performing perceptually guided actions on objects is representing the spatial location of this object in one's egocentric space: as in front of us, or to our left, etc. If we couldn't represent the spatial location of an object in our egocentric space, then we would have no idea which direction to reach out to grab it or use it for any other action. But, crucially, depicted objects are not represented in our egocentric space: the depicted space is not our egocentric space. And while we may represent the depicted objects as having a spatial location in the depicted space, we can't represent them as having a spatial location in our egocentric space" (2015: 189). Also, according to Nanay, "there is no fact of the matter about the distance between the perceiver and the depicted object. If I see a picture of an apple, there is always a fact of the matter about how far away the surface of the picture is from me, but there is no fact of the matter about how far away the depicted apple is from me. It is not represented as having an egocentric spatial location—thus, it is not represented dorsally" (Ibid.). This clarification about (b) is a further good reason to argue that our accounts are not in conflict, although I do argue about the notion that depicted objects elicit motor responses. I will come back to this in (§ 4.6).

So, depending on how we understand Nanay's claim – either the one reported in (2011a), or the one reported in (2015) - my personal claim has a different impact on the

---

[47] However, Matthen and Nanay share virtually the same ideas, albeit with some differences (Nanay 2011a: 477). For the purposes of my chapter, the important one is the exception for the necessity of (b). Moreover, in place of ventral vision and dorsal vision Matthen talks about, respectively, descriptive vision and motion-guiding vision. Here this distinction is not important.

[48] …whereas, as we have said, Matthen does, but given what I have clarified above, there is no real conflict.

[49] The same applies to the ventral representation of the surface (Nanay 2011a, 2015).

literature. If we properly address Nanay's account, in the light of the *précis* concerning (b) (2015), it is easy to understand that there is no conflict between Nanay's claim and mine, which is simply a neurological extension of his. However, if we understand this claim as defending that (b) is necessary, which is the claim originally defended by Matthen, instead my claim might seem to contradict this thesis. But I'll show that this is not the case in (§ 4.6).

Beside Nanay's and Matthen's proposals there is an interesting third proposal concerning the dorsal/ventral account of picture perception. The proposal is by Voltolini (2013), who is the only one to argue that the cases of eminegligent patients (and similar brain damaged people) who unconsciously mistook pictures as their subjects, suggest that, not only does the identificative component of their unconscious perceptual states appear to be directed towards such subjects rather than to the pictures' vehicles – cfr. with the four cases above proposed - but also the motor-guiding component is directed towards the depicted object. This is so because, as reported by Voltolini, there is evidence that eminegligent patients react to the grouping properties of what they face - even if they are not aware of this - both when they face a two-dimensional geometrical figure that has no pictorial value at all, and when they face a two-dimensional figure whose pictorial value can be retrieved. In the former case, they react to mere two-dimensional geometrical figures involving an optical illusion (e.g. Judd illusion), like normal individuals (they bisecate the figure more on its left-hand side than at its real center) even though unlike normal individuals they are not aware of facing an illusory figure. In the latter case, these patients react to two-dimensional geometrical figures involving Kanizsa-items (they see a Kanizsa-item (e.g. a square) within the figure even when its contours are not physically traced) like normal individuals, although unlike the normal individuals they are not aware of facing one such figure. So, they have an unconscious 'ventral' identification of what amounts to a picture's vehicle and they can also grasp a figure-ground segmentation like normal individuals: they can locate the Kanizsa item in front of a background, as if such an item were precisely a three-dimensional entity. They unconsciously grasp the pictorial value of the two-dimensional geometrical figure they face: they unconsciously see in the figure precisely a three-dimensional scenery (the Kanizsa item in front of a background). In other words, they have an unconscious 'ventral' identification also of what amounts to a picture's subject. However, in general, they manage to unconsciously identify both the picture's vehicle—the two-dimensional geometrical figure they face—and the picture's subject—the Kanizsa item (three-dimensionally conceived) in front of a background. This evidence is closely related to what I am arguing here, suggesting that sometimes the motor-guiding component of our brain can be directed at the depicted objects (cfr. with my § 4.6). However, in this chapter, I leave aside Voltolini's account, which is not devoted to the solution of the issue I am focusing on here. Moreover, I will use different evidence, which this debate on picture perception does not mention.

So this chapter deals with claim (b) defended in the literature: that the depicted object is not represented by dorsal perception, because dorsal perception localizes objects in egocentric action space. Egocentric space[50] does not overlap with the space represented by/in the picture. Therefore we cannot localize the depicted objects in our egocentric space, hence the depicted object is not represented by dorsal perception (Nanay 2011a). For this reason, only the encoding in egocentric coordinates – subserved

---

[50] I do not address the different notions of egocentric space. Here, egocentric space is action space, sometimes referred to in the empirical literature as the peripersonal space "within our reach". Moreover, Nanay appeals to a weaker notion of egocentric localization with respect to Matthen. For the purposes of this chapter this is not relevant. Indeed, even though with some differences Matthen and Nanay share almost the same ideas (Nanay 2011a: 477).

by the dorsal visual system - can detect the presence of real things offering action possibilities. But our dorsal stream attributes action properties to egocentric localized objects, which cannot be those in pictures, since egocentric action space does not overlap with the space represented by the picture. Therefore, we do not represent depicted objects as offering action possibilities, and, in turn, we cannot attribute action properties to the depicted objects (Matthen 2005)[51]. Moreover, since empirical evidence has shown that the dorsal subsystem is divided into two sides: the ventro-dorsal (V-D) and dorso-dorsal (D-D) subsystem, one responsible for manipulating objects (d-d) and one responsible for localizing in egocentric space (V-D) (Rizzolatti and Matelli 2003), it has been argued that we do not localize depicted objects in our egocentric space insofar as our V-D visual subsystem does not represent the depicted object (Nanay 2011a). I'll come back to this in (§ 4.5).

Now let us look into my proposal more precisely.

4.2 The Structure of my Proposal

Given the background ideas of the dorsal/ventral accounts of picture perception, it is now easier to understand what I mean when I say that the main claim of the present chapter is that both with real object perception and with the perception of depicted objects, our dorsal activity is activated, insofar as a portion of our dorsal stream, the ventro-dorsal stream, can represent depicted objects as having action properties. This claim constitutes one of the answers to (1), which were considered impossible by the other accounts.

As I did in (§ 3), before starting with my argument it is necessary to address the difference between (1) and a very different question

(2) what properties does our visual perceptual system respond to/covary with respect to the perception of depicted objects?

As I said in (§ 3). (1) is about what properties are attributed by the perceptual system (or what it represents entities as having) and not about what properties are out there (presumably in a causal relation with our perceptual system), as in the case of (2). The properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b; Burge 2010). I recalled this important philosophical point because here I will refer to it in a different manner with respect to (§ 3). Indeed, I will show (§ 4.3) that the answer to (2) is fundamental to reaching the answer to (1). I'll first show that it is so in the case of normal objects and, then, I'll show the similarity with the case of depicted objects.

The argument proceeds along the following lines. In (§ 4.3) I introduce evidence on a particular bifurcation of the dorsal stream, the ventro-dorsal stream. Saying that something is encoded by the ventro-dorsal stream implicitly means that the information has started to be mapped from the primary visual cortex, passes through the parietal cortex and arrives until the premotor cortex. Moreover, the ventro-dorsal stream contains a cortical circuit, the parieto-premotor network AIP-F5, which is the cortical circuit most involved in detecting action properties and in the construction of the motor acts in order to interact with those action properties. Here, I show that, in the case of normal objects, the ventro-dorsal circuit covaries with particular geometrical features of

---

[51] Matthen uses the term affordance. For philosophical reasons, I prefer to use here the expression action possibility, and, in turn, the related component of action possibilities called action property; see footnote 4.

the objects we are faced with - the answer to (2) - and that, in turn, it attributes action properties to the objects we are faced with – the answer to (1).

In (§ 4.4) I claim that our dorsal visual system functions in a somewhat similar manner when we are perceiving depicted objects. This claim is empirically well founded with evidence concerning the sensorimotor activation of the dorsal stream, the sub-components of the ventro-dorsal stream, as the parieto-premotor network AIP-F5, and, in general, the ventral premotor cortex, during the perception of depicted objects. This implicitly suggests that dorsal perception is active for the depicted objects. Hence, this can be considered as one of the true answers to (1). I also point out that, as in the case with normal objects, the ventro-dorsal circuit covaries with particular geometrical features which the depicted object seems to present – the answer to (2) – and that, in turn, it attributes action properties, on the basis of the translation of the particular geometrical features the depicted objects we are faced with seem to present. Thus, we can answer to (1) that our (dorsal) visual system attributes action properties to the depicted object. Remember that, as said before, the answer to (2) is fundamental to get the answer to (1). I also suggest that the functional interplay of AIP and F5 represents a possible explanation about dorsal response to depicted objects - as said in chapter 2 (see footnote 19) all the evidence I report here is grounded on studies on both human and non-human primates and an important overlap between the different cortical areas is widely agreed (see Borghi and Riggio 2015: 3; Shikata 2003).

My argument may seem completely explicated here. But given the dependence of dorsal perception to peripersonal encoding and given that the dorsal/ventral account of picture perception has defended the impossibility of peripersonal encoding with depicted objects, as said in (§ 4.1), something seems to be missing. So, in (§ 4.5), I specify that dorsal ascription of action properties to depicted objects is possible when the image of the target is apparently located within the subject's peripersonal space, that is when the target is perceived as apparently reachable by the observer.

At this point, my argument would really be complete: I have shown that depicted objects - those apparently located in the peripersonal space of the observer - are dorsally represented as having action properties. However, in (§ 4.6) I offer an explanation of how it is possible by reporting evidence that dorsal perception functions in a somewhat similar manner when it is faced with a normal object, or a depicted one, because it cannot really distinguish between depicted and normal objects; thus, once an object, whether depicted or real, is perceived as being in the peripersonal space of the observer, the dorsal stream responds to the apparent motor possibility fostered by the distal target, regardless of the fact that the source is a depicted object, or a real object. Here, I also address the implications of my proposal for the dorsal/ventral account of picture perception.

Although there are good reasons to affirm that the evidence I report is a good example of perceptual representations, I do not defend this idea here (see Nanay 2013b; Zipoli Caiani 2013; Butterfill and Sinigaglia 2012; Jacob and Jeannerod 2003). Furthermore, we know that perception can be conscious or unsconscious, and in turn, investigated from both the personal and the subpersonal level of description. Given that perceiving something in a picture does not need to be a conscious experience here I am concerned only with the subpersonal/unsconscious level of description.[52] Finally, note

---

[52] This position is not different from the one endorsed in the dorsal/ventral account (Nanay 2011a: 461). I leave aside the issue concerning the cognitive penetrability of the visual streams.

that the empirical literature uses the term affordances. I do not use this term, see footnote 51.

So, let’s start with the explanation of how our visuomotor system works in the case of face-to-face perception.

4.3 Visuomotor Circuits for the Detection of Action Properties in Normal Objects

Before showing that our visuomotor circuits function in a somewhat similar manner both in the case of seeing face-to-face and seeing depicted objects, I first have to explain how those circuits work in the case of face-to-face visuomotor processing, so I will introduce here the main visuomotor circuit for the detection of action properties in objects, which, as said, is due to the visuomotor component of MRs (2.3.1). For ease of reference with respect to the functional similarity concerning the functioning of the visuomotor component in relation to both normal and depicted objects, I briefly recall here the evidence about this component and its neural correlates, so that the reader can bear in mind what are the neural cells involved in this process and can understand how similar is this brain functioning in the case of picture perception.

As said before, it has been famously shown that the dorsal stream is further divided into two: the dorso-dorsal stream, (D-D, related to the superior parietal lobule, SPL) and the ventro-dorsal stream, (V-D, related to the inferior parietal lobule, IPL) (Rizzolatti and Matelli 2003). Although the dorsal stream is crucial for the visual guidance of actions, the transformation of intrinsic object properties into motor acts relies on a well defined cortical network lying in between the parietal and the ventral premotor cortex (Gallese 2007), that is, the ventro-dorsal stream, whose main components for this task are the anterior intraparietal (AIP) area and F5 (in the most rostral part of the ventral premotor cortex; though the ventro-dorsal stream has other crucial components). AIP is one of the end-stage areas of the dorsal stream and few processing stages separate the AIP area from early visual areas. Empirical data (Castiello 2005; Sakata et al. 1995; Murata et al. 2000; Baumann et al. 2009; Srivastava et al. 2009; Theys et al. 2012 a, b 2013) suggest that, since many AIP neurons respond selectively to objects during both passive fixation and grasping[53], it is the AIP that extracts visual object information concerning action possibilities for grasping purposes and relays this information to neurons in the area F5, with which it is directly connected (Borra et al. 2008), which then activate the primary motor cortex (for a review see Romero et al. 2014). Accordingly, evidence shows the presence of visuomotor[54] (canonical) neurons in F5, which encode the information received by AIP and translate the geometrical attributes of the objects into motor commands in order to interact with them. Note that also canonical neurons respond during object fixation, regardless of the actual execution of an action[55] (Murata et al. 1997; Sakata et al. 1995; for a review see Raos et al. 2006). Furthermore, the entire inferior intraparietal lobule (IPL, related to the V-D) — of which the AIP is a part —

---

[53] A large portion of neurons in this area discharge during object fixation and are selective for object properties, such as shape, size, and orientation (Verhoef et al. 2010).

[54] These “visuomotor” neurons showed a specific selectivity, discharging more strongly during the fixation of certain solids as opposed to others, the difference between them depending on the kind of grip afforded by those objects (e.g. precision grip, finger prehension, etc.).

[55] In experiments with monkeys, just as the subject looks at the object its neurons fire, activating the motor program that *would be* involved were the observer actively interacting with the object. The evoked motor potential (linked to MEP) is just a potential act. This will be important for the evidence in (§ 4.4).

receives direct inputs from the visual areas MT/V5, which in turn receive inputs from the primary visual cortex V1 and, as I have said, V-D (related to the IPL) projects to the areas of the ventral premotor cortex (areas F4 and F5). Note that V-D is the main site of convergence of the pathways from the extrastriate visual areas of the dorsal stream and this evidence suggests a direct connection between the visual areas that receive information from the retina and the sensorimotor areas of the parieto-frontal circuit (AIP and the premotor cortex), also called the ventro-dorsal stream (V-D) involved in visuomotor transformation. This evidence seems to show that on seeing an object we are at the same time getting its visuomotor priming (i.e. the visuomotor representation of its "action property"), and the internal *simulation*[56] of one of the actions we could perform upon it (i.e. the most suitable[57] motor program required to interact with it with respect to the action possibility), regardless of the actual action execution (Gallese 2000; Jeannerod 2006; Jacob, Jeannerod 2003). In a nutshell, when I am looking at the cup of coffee on my desk, the ventro-dorsal circuit is crucial for my purpose of grasping the handle of the cup. Consider that, first of all, the AIP detects the geometrical features of the handle that exhibit precise motor characteristics with respect to my motor repertoire. This means that the shape, texture, size are encoded. This information is then sent to F5, which, given the information received by the AIP, computes the most suitable motor act (say, a power grip) in order to grasp the handle of the cup. This is of course an over exemplification of the extremely complex functions of the ventro-dorsal stream. However, while it is widely agreed that the process can be divided into the two steps I have described, new evidence will give a more detailed account of the apparent fuzziness concerning the fluid connectivity of this extremely complex neural circuit. For example, it has been argued that the task of AIP neurons is simpler than grasping action possibilities with respect to the motor acts, and there's room to suppose that they encode action properties (the transformation of object geometrical features into action properties) and additional processing to encode the action possibilities in relation to the suitable motor acts requested by the situation, which seems to be due to F5, is necessary (Romero et al. 2014; Pani et al. 2014)[58]. However, for the purposes of this chapter the functional discrimination I have reported should suffice. We have only to keep in mind that action properties are those encoded by the process of visuomotor transformation, which, at its first step, reads object geometrical features as action properties. Action – or,

---

[56] There are different notions of simulation: here I mean an automatic mechanism with the perceptual function to facilitate the motor preparation (Gallese 2000, 2007, 2009). That is, when a visual stimulus is presented, it directly evokes the simulation of the congruent motor schema which, regardless of whether the action is executed or not, maps the stimulus position in motor terms. This is really important insofar as motor activation frames the represented action within the constraints of a real action: represented actions correspond to covert actions as neurophysiological simulation of the mechanisms normally involved in the physical action generation (Jeannerod 2006: 130-131). Also, motor activation is highly specific to the action that is represented. However, neural commands for muscular contractions are effectively present, but simultaneously blocked by inhibitory mechanisms (Jeannerod 2006: 2.3.3). From a philosophical point of view, this is really important because, while overt action execution is necessarily preceded by its covert representation and simulation, covert representation and simulation are not necessarily followed by overt execution (Jeannerod 2006: 2; chap. 2, 6). An important point: visuomotor representations deal with the process of transformation of the visual percept, that is, with the transformation of the geometrical properties of the objects into action properties and then into motor acts; the recruitment of the suitable motor acts, as a result of the visuomotor transformation, is linked to the motor simulation concerning the selection of a motor plan on the basis of the translation of the properties of the perceived object into action properties (Borghi and Cimatti 2010; Decety and Grezes 2006; Borghi et al. 2010) – see my (§ 2.3.1).

[57] The concept of "most suitable" is not linked to the demand character of action properties, but to the embodied motor possibilities of the subject with respect to the action he can perform. That is, the object is not prescribing a precise kind of action, even though our motor system detects only a small set of action possibilities to the extent that only few motor acts are possible due to bodily constraints.

[58] See footnote 60.

motor – possibilities are those encoded by the process of motor preparation, or simulation due to motor responses concerning the motor commands during object fixation with respect to the action properties encoded. Remember that these processes are activated also in the case of passive object fixation, regardless of the actual execution.

Now, coming back to what I said before, it is important to distinguish question
(1) What kinds of properties does our visual system attribute to the depicted object?

from a very different question

(2) what properties does our perceptual system respond to/covary with during pictorial perception?

As I said, (1) is about what properties are attributed by the perceptual system (or what it represents entities as having) and not about what properties are out there (presumably in a causal relation with our perceptual system), as in the case of (2). The properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b; see Burge 2010). I also said that the answer to (2) is fundamental to reaching the answer to (1). Well, according to the empirical evidence, the neural phenomena above described translate object attributes into motor commands: precisely, they represent S (geometrical objects features) as being F (action properties) and F (action properties) as Q (motor acts) resonating with the simulation of the performable motor act. There could also be room to suggest (see above the analysis offered by Romero et al. 2014) that the AIP reads geometrical features as action properties and relays this information to F5, which transforms the action properties into suitable motor acts, even though, in order to defend my claim, it sufficient to say that the ventro-dorsal stream has a critical role in the extraction of visuomotor properties of objects (Shikata et al. 2003). But I have said that the properties our perceptual system responds to or tracks may not be the same as the ones it represents objects as having (Nanay 2011b). This is to say that, concerning the empirical evidence, according to the visuomotor transformation process with normal objects, our perceptual system responds to/covaries with particular "object geometrical features". Accordingly, thanks to the visuomotor transformation, our visual system attributes "action properties" to the object. Thus, the properties our perceptual system responds to or tracks (geometrical properties) are not the same as the ones it represents objects as having (action properties). Yet, the detection of particular geometrical features allows these geometrical features to be transformed (or say, read) into action properties, or, in other words, to attribute action properties to the object. Furthermore, this encoding is accompanied by the motor resonance of both the two main components in the covariation and attribution: AIP-F5. Thanks to this, the motor resonance exhibited by the motor simulation is possible.

Here, I have explained how our visuomotor system works during face-to-face perception. In the next section, I show that our visual system functions in a somewhat similar manner when we perceive depicted objects. In order to defend this idea, I will introduce evidence concerning sensorimotor activation of the same areas I have talked about in this section, during the fixation of depicted objects.

4.4 Visuomotor Circuits and Depicted Objects

In (§ 4.3) I explained how those visuomotor circuits work in the case of face-to-face visuomotor processing. In particular, we saw that different portions of a chunk of the dorsal stream, the ventro-dorsal stream – the components of the parieto-premotor network AIP-F5, projecting to the ventral premotor cortex – which are the main cortical areas in visuomotor transformation, are activated both during passive fixation of objects and during action execution. Now I need to report evidence showing that these areas are also activated during the perception of depicted objects. In other words, visuomotor processing activated in face-to-face perception is activated in a somewhat similar manner when we are faced with depicted objects. This will be a way to show that depicted objects are dorsally perceived and that, in turn, we represent them as having action properties for which our motor preparation, or simulation, is highly selective and indeed activated.

Several empirical data show that visually presented images of objects activate the same sensorimotor areas functionally involved in the execution of the actions afforded by the depicted objects "in the flesh". In an experiment by Chao and Martin (2000), concerning the "representation of manipulable man-made objects in the dorsal stream", the authors have shown the presence of a deep selectivity of both the left posterior parietal and left ventral premotor cortex, in particular F5 canonical neurons, for pictures of tools, in comparison with different categories of objects. This suggests that, the encoding of visuomotor information in the visual stimulus, with the related representation of action properties based on the information of visuomotor nature (see § 4.3) is possible in the case of depicted objects as well and depends on the same neural underpinnings at the basis of the visuomotor transformations for overt motor interactions (see § 4.3). A similar conclusion can be drawn from the evidence by Proverbio et al. (2011), who showed that action possibilities (which in the experimental settings they call "affordances") are encoded - within the 250 ms time course – when depicted objects are presented; precisely, images of manipulable objects. Another interesting experimental (TMS) study comes from Buccino et al. (2009), who recorded the presence of motor-evoked potentials (MEPs) of parts of the hand involved in the motor act of grasping in subjects faced with images of damaged and undamaged objects, e.g. mugs with a whole or a broken handle (see also below) (this practice is well-known in neuroscience, see Anelli, Borghi and Nicoletti 2012). Buccino et al. also suggest the high involvement of dorsal perception in object knowledge during this task. A further example comes from Costantini et al. (2010) and confirms that the presentation of virtual images of objects elicits motor-related effects related to the action properties of the presented objects. Importantly, Grezes and Decety (2002) have shown the precise response, among other areas, of the IPL, that is, the ventro-dorsal stream. Finally, Borghi et al. (2007) showed that pre-motor areas are activated during the naming and viewing of pictures of manipulable objects (as in the case of Chao and Martin 2000; see also Gerlach et al. 2002; Craighero et al. 2002).

In a nutshell, while evidence reported in (§ 4.3) suggested that end-stage areas of the ventro-dorsal chunk of the dorsal stream, such as the ventral-premotor cortex, in particular its canonical neurons, are active even during fixation of normal objects, that is, without execution, evidence here suggests that, similarly, the activation of those areas is possible in the perception of depicted objects, even if there is no request to perform any kind of motor interaction upon the visually presented target, insofar as the task consisted only in merely observing the depicted objects on the screen, or in photographs (with different requests on the subjects, and different experimental settings Chao and Martin 2000; Grezes and Decety 2002; Buccino et al. 2009; Cardellicchio et al. 2011; Costantini et al. 2010; Tucker and Ellis 1998; for an excellent review concerning picture perception with respect to the activity of the (pre)motor areas see Zipoli Caiani 2013)[59]. Now I will

---

[59] See the case of canonical neurons in (§§ 4.3, 2.3.1).

report evidence that the same resonance with pictures is also given by the AIP.

Evidence by Romero et al. (2014; see also Pani et al. 2014), show that a subset of AIP neurons is also activated by two-dimensional images of objects and even by outline contours defining the object shape: neurons in the posterior AIP (pAIP) are selective for two-dimensional (2D) images of objects, a selectivity that is primarily based on the contours of the object shape (Romero et al., 2012), and this is fundamental in order to understand the sensorimotor resonance in the case of pictures. Note that neurons in the anterior AIP (aAIP) area also respond to images of objects, though are less frequently selected. However, the search stimulus set consisted of 21 2D area-equalized static images of objects (faces, hands, fruits, branches, and artificial graspable objects) termed "surfaces" because they all contained texture, perspective, and shading information. The authors recorded the activity of 82 single AIP neurons that responded selectively to images of objects during passive fixation. Neurons active during grasping were then tested in the passive fixation search test. All neurons responsive during grasping were also responsive during the passive fixation search test. The strong responses they observed confirmed that the recordings were performed in the AIP. Moreover, given the position of the AIP area in the hierarchy of visual areas, their finding that the neural coding of object shape in the AIP is largely based on a selectivity for small line fragments, which in their test appear to be more suitable for activating neurons in early visual areas (V1 or V2), and given that other studies are consistent with this low-level visual selectivity (for a review see Romero et al. 2014) concerning very short response latencies in the AIP area, they suggested that few processing stages may separate the AIP area from early visual areas. Given these anatomical and functional characteristics of AIP, the interplay between the AIP and F5 is crucial in the explanation of how our sensorimotor – in particular, visuomotor - areas can covary with particular (2D) geometrical properties which in turn can be represented as action properties – or, in other words, on the basis of the transformation of which (the geometrical properties) our visual system attributes action properties to the depicted object – (and this is possible due to the visuomotor mechanism I described in (§ 4.3) concerning real objects perception).

It is worth noting that the selectivity of the AIP for two-dimensional stimuli - and its interconnection to F5 - reported by Romero could explain the passage from the detection of a two-dimensional stimulus to the reconstruction of the suitable motor act to perform: the reconstruction of the possible motor interaction starts from the two-dimensional features of the object and continues regardless of the fact that the features pertain to a real object or to a depicted object. The covariation of the same geometrical properties that an object in perspective in a picture and a real object in perspective share leads to the same representation of action properties, and this is possible because the dorsal system doesn't distinguish between the distal cause encoded (cfr. with § 4.6 below). Indeed, the same action possibility encoding holds both when the perceiver is in front of a real object, and when the perceiver is in front of pictures of it, showing that the perceptual system is attuned to shared sensorimotor layouts and configurations in both the real and the depicted target, without distinction from the distal source[60].

Given the presence of this same mechanism with respect to (1) and (2) both for real objects that for depicted objects, it is possible to claim that action properties are

---

[60] For technical, neurological details concerning AIP encoding with respect to action possibilities and geometrical figures and F5 encoding with respect to motor acts as well as the different contribution in the encoding on those by AIP and F5 see (Romero et al 2014).

dorsally represented both in the case of perception of real objects perception and in the case of the perception of depicted objects.

This section reported sufficient evidence in order to suggest that dorsal representation of action properties is possible for depicted objects. But, we know that dorsal processing is linked to the encoding in peripersonal coordinates. So, one might wonder about the possibility of having a dorsal representation of action properties in the perception of depicted objects, even though it has been denied, by the dorsal/ventral account of picture perception, that they can be egocentrically encoded as much as normal objects can. In the next section, I address empirical evidence showing that different experimental settings can lead the subject (and her/his dorsal vision, see below § 4.6) to apparently perceive the depicted object as falling within her/his peripersonal space and as apparently reachable. In these cases, dorsal perception can attribute action properties to the depicted objects. I also report the neuroanatomical underpinnings of this process.

4.5 Depicted Objects and Action Space

In the footnote 50 I said that I use the expressions "egocentric space" and "peripersonal space" interchangeably and as near synonyms in denoting the action space, that is, the space within our reach (for a review concerning this synonymy see Holmes and Spence 2004). This space is linked to the motor perspective as well as to the motor capabilities/skills of the individual agent. Here, the reference to the anatomo-functional relation between the visuomotor component and its neural correlates (§ 2.3.1) and the egocentric/peripersonal component, and its neural correlates is crucial (§ 2.3.3) insofar as I am going to suggest that this functional link is maintained even in the case of the perceptual representations of depicted objects.

While the main purpose of this chapter is just to show that dorsal perception can ascribe action properties to depicted objects, an important point needs to be addressed here: results from motor neuroscience suggest that motor resonance is often[61] due to peripersonal/egocentric encoding of the target. The same holds for depicted objects as well - even if, the fact that the depicted object is perceived as falling in the action space and as being reachable is, of course, only apparent, and it is fostered by the experimental settings. For example, the experiment by Costantini et al. (2010) clearly shows that the motor response with respect to action possibilities is deeply dependent on the perceived spatial location of the object and on the apparent possibility[62] of interaction for the subject. Costantini and his colleagues presented their subjects with virtual images of a mug and divided the space into both reachable and non-reachable subspaces by presenting the image of the mug either in front of or behind a transparent panel. The compatibility effect was detected only when the image of the target was perceived as located within the subject's peripersonal space[63]; that is, when it was apparently reachable by the observer. This result suggests that the motor response is due to the peripersonal encoding of the depicted object, which is apparently presented in the peripersonal space of the observer and is apparently reachable. In accordance with this

---

[61] For exceptions, see footnote 67.

[62] The notion of apparent possibility of interaction will be important given the nature of dorsal processing. I will talk about this in (§§ 4.5, 4.6).

[63] As already said, for the purposes of this chapter, it not necessary to question the concept of egocentric space. Here I use this as a synonym of peripersonal space: both are concerned with action space, or the space within one's reach.

evidence, Janyan and Slavcheva (2012) tested the encoding of action possibilities using symbols superimposed on a picture of a pan, showing a faster response when the visual action possibility of a graspable object (e.g., a pan) corresponds to the response location and thus suggesting that the location[64] of the virtual depicted object is important for inducing motor resonance. In other words, the activation of the visuomotor response is modulated by a non-trivial egocentric relation with the depicted target due to the apparent possibility of interacting with the target's functional parts (e.g., a handle), as well as by their apparent spatial location: once again, response is found when the depicted object is apparently located within the subject's peripersonal space, and when it is apparently reachable by the observer (see Zipoli Caiani 2013 for this point).

As just said in (§ 2.3.3), the importance of peripersonal encoding of the target for detecting its action properties is widely agreed upon in the literature (De Stefani et al. 2014; Ambrosini and Costantini 2013). For example, F5 canonical neurons in the ventral premotor cortex selectively respond only to those objects presented in the peripersonal space but not in the extrapersonal space (Bonini et al. 2014). This spatial constraint holds for the general activity of the ventral premotor cortex as well (Maranesi et al. 2014; Turella and Lignau 2014; Bonini et al. 2014). More generally, action possibilities are detected, and the related motor acts are encoded, only when objects fall into the action space (Borghi and Riggio 2015: 7; Fattori et al., 2001, 2005; Hadjidimitrakis et al., 2011, 2013; Bosco et al. 2014).

This is due to corto-cortical functional interconnections of different areas within the circuits of grasping (Castiello 2005, Turella and Lignau 2014), precisely, the two main parieto-premotor networks the ventro-dorsal stream can be divided into: AIP-F5 and VIP-F4 (Rizzolatti and Matelli 2003; Gallese 2007). The circuit crucially involved in the mapping of our peripersonal space is constituted by the ventral intraparietal area (VIP) and F4. F4 occupies the posterior sector of the ventral premotor cortex, next to F5, with which it is adjacent. The VIP is located in the fundus of the intraparietal sulcus (Gallese 2007, Rizzolatti and Matelli 2003). AIP-F5 and VIP-F4 circuits are extremely interconnected from an anatomo-functional point of view: largely segregated parietofrontal connections link the rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4) (Luppino et al. 1999). Indeed, the VIP-F4 circuit codes the peripersonal space, which coincides with the motor space for arm reaching (Borghi and Riggio 2015: 9)[65]. Thus, canonical neurons space dependency is due to the connections between area F5 and area F4 (Fogassi et al. 1996; Matelli et al. 1996; Borghi and Riggio 2015: 7; Costantini et al. 2011a, b)[66]. Accordingly, motor simulation is spatially dependent as well: only those objects presented in the peripersonal space are automatically mapped, from a motoric point of view, through simulation (ter Horst et al. 2011, Gallese 2005, Grade et al. 2015).

My claim was that dorsal perception can ascribe action properties to depicted objects. Following this neural evidence, I can suggest that this is likely to happen when the image of the target is apparently located within the subject's peripersonal space, and

---

[64] See also Vainio et al. (2011), who revealed a compatibility effect for the orientation of the image of a graspable object.

[65] As in the case of the connection between the AIP and F5, and VIP and F4, the same extremely complex interplay holds between the entire parieto-premotor networks VIP-F4 and AIP-F5 (see Rizzolatti and Sinigaglia 2008).

[66] See footnote 67.

when it is apparently reachable by the observer[67] (the object in the monitor is perceived as being located within the peripersonal space of the onlooker). Indeed, crucially, this is in line with the different experimental settings in which presented images (Chao ad Martin 2000), or transparencies (Ellis and Tucker 1998) are presented near to the participants' heads 45 cm in front of the screen (Ellis and Tucker 1998), or 57 cm, (Cardellicchio et al. 2001, but see also Romero et al. 2012, 2014), that is, in which the depicted objects are apparently presented within the peripersonal space of the observer and are apparently reachable. I will consider the implications for the dorsal/ventral account below (§ 4.6), where I will report further evidence in line with these reported here.

I want to point out that my claim that these objects are represented as apparently being in the peripersonal space of the observer and apparently reachable is not problematic in the literature (see Zipoli Caiani 2013). In (§ 4.6) I explain why such apparent, but non-trivial, egocentric relation with the depicted object is possible for dorsal subpersonal vision, that is I provide further technical details about how dorsal subpersonal vision represents depicted objects as apparently located in the peripersonal space of the onlooker and attributes action properties to them and I also explain that my

---

[67] There is also evidence that sometimes egocentric/peripersonal encoding might not be necessary. Cardellicchio et al. (2011) showed that, whereas MEP can be modulated by the orientation of the part of an object which is suitable for motor interaction, it can also be modulated by its spatial location, regardless of the fact that this location overlaps with the peripersonal space of the observer. Similarly, Tucker and Ellis (2001) suggested that in order to detect action possibilities it is not necessary for the object to be placed in the peripersonal space. This appears to conflict with the result by Costantini et al. 2010. However, as Zipoli Caiani (2013: footnote 5) observes, while in the experiment by Costantini et al. 2010, motor elicitation is observed only when the image of the target was apparently located within the subject's peripersonal space and apparently reachable by the observer, contrarily, for Tucker and Ellis (2001), "affordance-related compatibility" effects (as they are called in neuroscience) are not dependent on the object being presented within the actual reaching space of the observer. This is because in the experimental settings of Tucker and Ellis, objects are presented behind a plastic LCD screen that inevitably biased any attempt to modulate their reachability by varying the spatial distance between the target and the observer (see also the end of § 4.6 for similar discussions). I would like to point out that, moreover, while in the case of Costantini et al. action is evoked "only" when the object is presented within the portion of the peripersonal space, in the case of Cardellicchio et al. (2011) also non-reachable objects can evoke motor responses; however, the authors found "higher" MEPs during the observation of graspable objects falling within the reachable space compared to the observation of either a non-graspable object or a graspable object falling outside the reachable space. But, as the authors affirm, this evidence clearly indicates that visuomotor recruitment is spatially constrained, as it depends on whether the object falls within the actual reaching space of the onlooker. Note that, in the case of Cardellicchio, we are still talking of virtual two-dimensional images of objects on a computer screen with a resolution of 1024 horizontal pixels by 768 vertical pixels, at a distance of approximately 57 cm (1370). The response may arguably be due to the fact that the computer screen may be perceived, by the subjects, as being in the peripersonal space, and thus, even those objects that are apparently located in the virtual extrapersonal space are perceived as being reachable - and, maybe, just smaller. On the other hand, in the case of Costantini, the presence of the panel influenced the visual perception of the participants, as in the case of Tucker and Ellis (2001). I am not concerned with this exception here. Indeed, my claim is that dorsal perception can ascribe action properties to those depicted objects which are apparently located within the peripersonal space of the observer and which are apparently reachable and this is sufficient in order to extend the dorsal/ventral account of picture perception. I just require that action properties can be represented on the basis of motor responses. Establishing whether or not this is really possible also in the case of those objects which cannot be egocentrically encoded goes beyond the purposes of this chapter. Finally, the difference in experimental setting may be investigated from the perspective of inflection. Inflection is the phenomenon according to which some of the seen features of the picture's surface are relevant in characterizing some features of the depicted object as seen in that vehicle (see Voltolini 2013, Nanay 2010b). Here I do not need to account for this phenomenon, insofar as different (similar kinds of) pictures induce the same "dorsal" response. Moreover, as I have said, I am not concerned with particular kinds of pictures involving *trompe l'oeil* - I'll be back to this issue in chapter 5.

idea is not in conflict with the dorsal/ventral account of picture perception. Also, while my claim is about the subpersonal level of perception, insofar as it is about dorsal vision, some authors have suggested that even our conscious ventral stream can quasi-egocentrically represent, to some extent, depicted objects as well (Briscoe 2009). Indeed, the idea that we can quasi-egocentrically perceptually represent depicted objects and their relative depth cues is an uncontroversial idea in both philosophy (for a brief review see Briscoe 2009: 447, 6.1–6.4; Grush 1998, 2000, 2007; see also Schwenkler 2014) and neuroscience (Pani et al. 2014; Bruzzo et al. 2008; for a complete review see Vishwanath 2014; Cutting 2003; Hecht et al. 2003): "perceptual discrimination of 3-D object and surface layout in films and photographs involves a quasi-egocentric perception of depicted space inasmuch as, when watching a film or looking at a photograph, one non-reflectively assumes the perspective of the camera" (Briscoe 2009: 447). As Briscoe observes, this idea is also compatible with the results of the TVSM – for example, with the evidence that grasping can be influenced by the size-contrast illusions (p. 447).

An important clarification here is that when saying depicted objects offer action possibilities this does not mean, of course, that we can act upon them. It just means that some part of our dorsal brain behaves – with respect to the neural firing - as if the action properties pertaining the depicted object, such as its geometrical configuration, pertained to a real 3D object. Indeed, the simulation linked to the motor act we can perform in the given motor situation - on the basis of the visuomotor transformation of object attributes into motor commands. The same visuomotor transformation, with the simulation of the motor act, is activated both in the case in which the geometrical features pertain to a real object and to a depicted object[68]. One could argue that this evidence shows only that the dorsal stream responds to pictures because it is involved in the perception of the surface of the pictures, that is (c) The picture surface is represented by dorsal perception. However, evidence shows that the affordance related effects[69] registered are deeply related with the kind of motor act (e.g. power grip) one can perform on the depicted object (e.g. the handle of a mug), which is, in these cases, different from the act one can perform on the picture surface (precision grip)[70]. In many of these cases, as we saw, pictures are presented on a monitor, which, of course, cannot afford the same action afforded by the depicted object. In other words, looking at an image of an object triggers the activation of a suitable motor pattern for the execution of actions and the motor activation is highly specific to the action that is represented (Jeannerod 2006, see my footnote 56). Once again, think about the case of Buccino et al. where subjects were asked to observe virtual images of objects, in this case of handles. Note that in this experiment the motor-evoked potentials (MEPs) are from the right opponent pollicis and from the first dorsal interosseous muscle. These anatomical components are crucial in grasping, and the presence of this kind of motor response shows us that the motor act encoded pertains to the handle and not, of course, to the surface of the image, since in this case the image is not a normal picture, but an image on a monitor, which cannot involve grasping. The experiment shows that motor response is dependent on particular pragmatic features of objects (for example, a handle) (see my footnote 56). Thus, here (c) is not the case. The same holds for the other similar experimental settings. Also, following what I said at the beginning of the section, while the subjects may not be able

---

[68] See footnote 56.

[69] Here I have maintained the neuroscientific terminology.

[70] The practice of using pictures in cognitive neurosciences to study how seeing tools automatically activates motor information. I cannot survey all of them; see (Craighero et al. 2002; Ranzini et al. 2011; Borghi et al. 2012). See § 4.6.

to actually act on these action properties, nonetheless to be highly specific representations of action properties they should have non-trivial relation to egocentric representations of the depicted object, and this is what happens in the experimental setting in which the depicted objects are perceived as apparently reachable because apparently located in the action space of the observer. But I have a further empirical explanation for this non-trivial egocentric relation of dorsal vision with the depicted object in (§ 4.6).

Now, we can move on to the next section, in which I report evidence concerning the dorsal stream which might give a further empirical reason to the claim exposed in (§ 4.4) and, in turn, go towards the idea that dorsal perception can detect more of those common cues shared by face-to-face perception and picture perception than we can expect if we were following the dorsal/ventral account on pictures.

Before doing this, another important point needs to be addressed. As we have said, in claiming that dorsal vision cannot represent the depicted object, Nanay has also recalled evidence that the dorsal subsystem is divided into two sides: the ventro-dorsal (V-D) and dorso-dorsal (D-D) subsystem, one responsible for manipulating objects (D-D) and the other responsible for them in egocentric space (D-D) (Rizzolatti and Matelli 2003); hence, he has reformulated the claim (b): we do not localize depicted objects in our egocentric space, that is, our v-d visual subsystem does not represent the depicted object (Nanay 2011a)[71]. But I clearly showed that the ventro-dorsal stream can represent depicted objects which are perceived as being in the peripersonal space of the observer, as having action properties[72] - see also my (§ 2.3.3).

Furthermore, in accordance with the insights of Rizzolatti and Matelli (2003), the difference between D-D and v-d is not that manipulation is encoded in D-D while egocentric localization is encoded in V-D, because also V-D is concerned with manipulation; in fact, this is the main cortical portion involved in the visuomotor translation (§ 4.3). Following a distinction made between variable and stable mechanisms pertaining to the detection of action possibilities, it has been suggested that the former are performed in the D-D, while the latter in the V-D (this is not the place to focus on this topic; see Borghi and Riggio 2009, 2015; Sakreida et al. 2013): D-D is involved in the online interaction with objects (in new motor situations), while V-D is responsible for how our knowledge of objects influences the way in which we represent them.

So, the distinction Nanay is looking for can be found in those two different parieto-premotor networks in the V-D – AIP-F5 and VIP-F4 - and not between D-D and V-D: it is one portion of V-D, namely the parieto-premotor network VIP-F4, that is involved in the egocentric localization and it is the other portion of V-D, namely the parieto-premotor network AIP-F5, that translates those motor geometrical features of

---

[71] "My strategy was to show that we do not localize depicted objects in our egocentric space. Thus, the argument I presented in this section, rephrased using the terminology of the three visual subsystems framework, aimed to show that our ventro-dorsal visual subsystem does not represent the depicted object" (Nanay 2011a: 470). See also the notion of "the perceptual individuation of visual objects by location" in Jacob and Jeannerod (2003).

[72] Note that, while Nanay suggests that D-D deals with manipulation, whereas it is V-D that deals with egocentric localization, we have evidence that neuronal activity in area V6A of the dorso-dorsal stream can also specify object position with high specificity for the peripersonal (reachable) space not only during reaching tasks (Fattori et al. 2001, 2005; Hadjidimitrakis et al. 2013; Bosco et al. 2014). However, data suggest the leading role of V-D concerning reaching and grasping activities (Maranesi et al. 2014; Turella and Lignau 2014; Gallese 2007; Rizzolatti and Matelli 2003; Borghi and Riggio 2015; Bosco et al. 2014, Castiello 2005). For the purpose of this chapter, this should suffice.

objects into motor properties (and this is possible, both for picture perception and for real object perception) (§ 4.4). However, it is important to keep in mind that the neural correlates for visuomotor transformation and those for egocentric localization are really interconnected, so that talking about a possible divide concerning their processing does not make much sense[73] - for a review see (§ 2.3.3). I will come back to the implications of this remark in the dorsal/ventral account of picture perception at the end of (§ 4.6).

To summarize (§§ 4.4, 4.5) show something important. First, like normal objects, depicted objects can be dorsally perceived as well, since the visuomotor system is activated by the apparent possibility of interacting with the depicted object's functional parts exhibiting motor-related properties (e.g., a handle), even if, of course, they do not really afford the apparent motor interaction they seem to evoke (§ 4.4). Second, this happens when the subject (or, her/his dorsal vision, see below § 4.6) entertains a non-trivial peripersonal/egocentric relation with the depicted object, that is, when the depicted target is apparently presented, in the experimental setting, in the peripersonal space of the observer and perceived as apparently reachable, even if the task only consists in the observation of virtual images of objects.

So, all I have said till now has been sufficient in order to defend my claim that depicted objects can be dorsally represented by dorsal perception as having action properties. However, in the next section (§ 4.6), I would like to address further evidence related to the fact that dorsal perception entertains a non-trivial egocentric relation with the depicted target suggesting a possible philosophical explanation of the empirical reason for why those depicted objects we face and perceive as apparently located in the peripersonal space and as apparently reachable are able to foster motor-related effects in our dorsal vision. In turn, I will discuss the implications of this explanation for the dorsal/ventral account of picture perception, suggesting that there is no conflict with my claim.

4.6 Action, Pictures and the Dorsal Stream

So, now we know that dorsal perception can attribute action properties to these depicted objects which, in the experimental settings, are apparently presented in the peripersonal space of the observer and are perceived as apparently reachable. Here I report further evidence suggesting that I am on the right track in arguing that depicted objects - those apparently presented within the peripersonal space of the observer - especially of "3D" objects, such as tools - activate the same processing areas in the visual brain as are activated by those real objects of the same kind, and that this is the case not only for the ventral stream – this is claim (a) of the dorsal/ventral account (see Nanay 2011a, 2015) - but also for the dorsal one.

For example, neurons in the dorsal intraparietal sulcus selectively respond to depicted objects exhibiting particular depicted 3D shapes and orientations (Nelissen *et al.* 2009, James *et al.* 2002, Taira 2001, Sakata *et al.* 2003)[74]. And this is in line with the

---

[73] The same dependence between the two processings seems to hold for the D-D. See footnote 72.

[74] Tsutsui et al. (2002) explored the sensitivity of caudal intraparietal (CIP) neurons in the dorsal stream to texture-defined 3D surface orientation. CIP neurons are involved in high-level disparity processing (the reconstruction of 3D surface orientation through the computation of disparity gradients). Some CIP neurons are sensitive to texture gradients, which is one of the major monocular cues. Some of them are sensitive to disparity gradients, suggesting their involvement in the computation of 3D surface orientation. Moreover, those sensitive to multiple depth cues were widely distributed together with those sensitive to a specific depth cue, suggesting the involvement of CIP neurons in the integration of depth information from different sources. The convergence of multiple depth cues in CIP seems to play a critical role in 3D vision by constructing a generalized representation of the 3D surface geometry of objects (Tsutsui et al. 2005). See footnote 80. I'll be back to this in chapter 5.

evidence exposed above about both AIP and F5. Indeed, there is evidence that F5 is better suited for the decoding of the grip type, while AIP is more accurate in predicting object orientation (Baumann, Fluet and Scherberger 2009; Gallivan and Wood 2009; Fluet, Baumann and Scherberger 2010) – for an analysis concerning depth perception and 3D perception in relation to the AIP-F5 circuit see (Theys et al. 2015). In line with this, dorsal perception seems to be sensitive, to some extent, to depth cues apparently exhibited by depicted objects (Gonzalez *et al.* 2008; Berryhill and Olson 2009; Marotta and Goodale 2001) and seems to discriminate between images of depicted graspable and non-graspable objects (Rice et al. 2007). In accordance, the most important end-stage of the dorsal stream concerning visuomotor interaction, the ventral premotor cortex, is highly selective during manipulable objects observation with respect to non-manipulable objects: images of tools, but not of houses, animals, and faces, activated the ventral premotor cortex (Grafton et al., 1997; Knight et al., 1999; as in the case of Chao and Martin 2000), and that the ventral premotor cortex was activated with manipulable objects but not with non-manipulable ones (e.g., Gerlach et al. 2002; Kellenbach et al. 2003) (this experimental evidence is crucial for what I said in § 4.4). This is sufficient for my claim (for the different activations of dorsal and ventral perception with different images see Proverbio et al. 2011).

However, notwithstanding these characteristics of the dorsal stream, there is one feature that seems to be crucial for the claim defended in this chapter: dorsal perception does not discriminate in a precise manner between objects seen face-to-face and depicted objects, because this capacity is subserved by the ventral stream[75]. Indeed, as shown by (Westwood et al. 2002), the dorsal stream does not construct a *complete* 3D structural description of the target object. Such a description, nonetheless, is necessary for response *selection*, in order to detect the action afforded by an object, or in the case of pictures, to understand that there is no possible interaction[76]. Accordingly, it is the ventral stream that plays the key role in object recognition and response selection, distinguishing between 3D objects and 2D images of objects by detecting conflicts between various visual cues[77] and selecting different visuomotor strategies for a 2D image versus a 3D object[78]. Thus, ventral perception selects an appropriate response for a particular visual stimulus, based on a comprehensive analysis of object volumetric structure, whereas the dorsal stream plans the precise metrics of the intended action, based on a pragmatic analysis of the object's spatial features. In other words, dorsal perception can control grasping without ever constructing a volumetric representation of the distal stimulus, which is due to the ventral stream. However, though dorsal perception explores depth cues and 3D shape perception without constructing a precise volumetric representation of the target object, which is due to ventral perception[79], we saw in (§ 4.4) that 2D structure is sufficient in order to activate the dorsal detection of action properties[80].

---

[75] Accordingly, prior to discriminating depicted objects as such, infants seem to perceive depicted objects as real objects affording action and they even grasp at the pictures as if trying to pick up the depicted objects (DeLoache et al. 2004: 68; see also Pierroutsakos and DeLoache 2003; Deloache et al. 1998).

[76] Volumetric object representation is necessary for the visual control of grip formation and response selection, to ensure that we do not attempt to reach for objects that cannot be grasped (Westwood et al. 2002).

[77] It is computationally efficient for one visual system to handle both response selection and object recognition: both require complete/detailed information about 3D object structure (Goodale and Milner 1992).

[78] One must perceive objects to be different in order to treat them differently.

[79] For the difference between shape perception and volumetric object recognition see (Briscoe 2008).

[80] We have further evidence that picture perception and face-to-face perception are, in general, very close with regards cues relevant for action. Indeed, while the perception of depicted objects lacks a precise sensation of a possibility of complete egocentric localization (in line with the claim of Matthen and Nanay)

This evidence might offer an explanation for why dorsal perception can represent depicted objects – at least those apparently presented within the peripersonal space of the observer - as having action properties.

We know that dorsal perception ascribes action properties to both normal objects (§ 4.3) and depicted objects (§ 4.4); this representation seems to be possible when normal objects are presented in the peripersonal space of the observer, or when, in the experimental settings, depicted objects are apparently presented in the peripersonal space of the observer (§ 4.5). So, the basic point is that dorsal perception entertains a non-trivial egocentric relation with the target, regardless of the fact that the object is real or depicted. However, dorsal perception distinguishes between graspable objects - like tools - and non-graspable objects, but not between real and depicted objects. This means that it attributes action properties to the object which is presented in the peripersonal space - or depicted objects apparently presented in the peripersonal space – regardless of the nature of the distal target that is causing the selective motor resonance; that is, regardless of the fact that it is a depicted object or a real one, to the extent that, after all, it cannot distinguish between them - note also that, most of the time, these depicted tools are the same size as a normal tool (that is, scale one to one) we use in our everyday life, so that, for the dorsal subpersonal processing, it is very difficult to discern between a real tool and a picture of it, which, in the eyes of the subpersonal dorsal computations are very similar objects. In other words, since dorsal vision cannot distinguish between real and depicted objects, everything that is apparently found in the peripersonal space can be represented as having action properties, fostering the apparent possibility of interaction: in order to foster visuomotor activation in dorsal vision, the important thing is to be faced with a geometrical configuration that is usually linked to an action property (see § 4.4), and that is apparently located in the peripersonal space of the observer, perceived as apparently reachable, no matter if the configuration pertains to a 2-D object or a real object[81].

Summing up, due to its particular computational nature, dorsal vision can entertain a non-trivial egocentric relation with depicted tools that are perceived as apparently located within the peripersonal space and as apparently reachable and can thus attribute action properties to them, insofar as, in the eyes of the subpersonal dorsal perception, there is a real egocentric relation with the depicted target.

All this suggests that the dorsal peripersonal/egocentric representation of the vehicle, which enables the consequent attribution of action properties to the vehicle is one thing, while the dorsal peripersonal/egocentric representation of the depicted object, which enables the consequent attribution of action properties to the depicted object - what I am concerned with in the chapter - is another.

All I said is strengthened by empirical evidence that motor response is activated when simple 3-D contours of the shape of the depicted objects, salient from a motor point of view are detected; for example, different neural populations of both AIP and F5 respond to the contours of the depicted objects which can be extracted as salient

---

that a normal object can offer, because the perception of depicted objects leaves absolute depth cues optically unspecified, depicted objects can foster in us the experience of 3D shapes, orientations, relative sizes, and depth ratios, which resemble those real properties found in ordinary 3D objects, for which egocentric localization is not needed (in accordance with § 4.4; for a review see Vishwanath 2014). This strengthens the fact that we can represent depicted objects as having action properties. Furthermore, if, according to recent evidence, we were to explore the connections between the two streams, we would have a more complete account of how different cues for action can be transferred between the streams. This would better explain the resemblance between face-to-face perception and picture perception concerning all the empirical data reported here on action properties. I do not address this issue here.

[81] This is in line with what I said about the AIP and its interplay with F5 during the perception of depicted objects (§ 4.4). I'll be back to this in chapter 5.

geometrical properties automatically represented as action properties. Indeed, AIP's selectivity is primarily based on the contours, aspect ratio and orientation of the depicted objects (Romero et al. 2013). It's not by chance that, in this circuit, 3D-shape selective neurons are co-localized with neurons showing motor-related activity, indicating, even in the case of depicted objects, a close relation of visual and motor information on the same clusters of neurons (Theys et al. 2015). This is in line with the evidence of several action-modulated perception effects that improve processing of action relevant visual features (Gutteling et al. 2015) and with the evidence that action preparation originating from activation in higher-order visuomotor control brain areas downstream from early visual cortex in the dorsal stream modulates visual processing according to the visuomotor constraints of the action that is encoded (van Elk et al. 2010; Gutteling et al. 2013). There are also other interesting aspects of dorsal processing in relation to 2-D perception. However, a focus on them would go beyond the scope of this chapter.

All this strengthens what I said about the unexpected functioning of MRs (cfr. with § 2.5): at the beginning I mentioned a widespread temptation, among those who use the dorsal/ventral account: assuming that, since we cannot really grasp a depicted object, because our visual system cannot egocentrically represent the depicted object, we cannot represent its action properties, which is the main task of our dorsal stream. I suggested that dorsal perception can in fact represent depicted objects as having action properties. However, I now need to address other common intuitions linked to this temptation, which I reject. First, that our visual conscious system – the ventral stream - detects the object of interest for us, establishes if it is actable upon - if it is a real 3-D object - and then sends the information to our visuomotor brain which in turn, only at this point, encodes a suitable motor act. This is because it is widely believed that it is visual consciousness that establishes whether the computation of the coordinates for motor action can start. This would depend on the result of the response selection that establishes whether we are faced with a reliable motor scenario (a real object) or not (a picture). In other words, it is common idea that, first of all, our conscious visual system detects a reliable motor situation, then, if this is the case, the information can be sent to the subpersonal visuomotor system, which can now compute the information necessary in order to act. This would suggest that the visuomotor system is not activated in the case of depicted objects, because visual consciousness would never send the command (with the relative) information for translating the properties of a depicted object into a motor act. After all, we cannot act upon a picture. But evidence shows that things are different. The visuomotor encoding is activated, despite the distal stimulus, despite the fact that – and even before – our visual conscious system has computed whether the object is "real" or not. Indeed, just a few indeterminate cues are sufficient in order to trigger the visuomotor transformation. In other words, before our conscious visual system encodes the appropriate response selection concerning a reliable motor scenario and establishes that we are dealing with a reliable motor situation (a real object and not a picture) the suitable motor act computed on the basis of the 2-D geometrical properties of the target is already stored in our motor quiver, regardless of whether the actual overt execution would be followed, that is, regardless of whether we can interact with the object or not. So, the visuomotor transformation is activated simply because our visuomotor system doesn't need any trigger, or any confirmation from our conscious visual system in order to start the visuomotor translation. Rather, the translation is already prepared and will be used at the discretion of the conscious response selection encoding the motor scenario. This is linked to the fact that, most of the time, we do not guide action with visual consciousness. Instead, our brain encodes our motor acts due to automatic processes, which can be then corrected or controlled[82].

To sum up, the evidence reported in (§ 4.6) suggests that our dorsal stream cannot discriminate between depicted and real objects. I have said that the dorsal stream is activated for similar cues that pertain to both face-to-face perception and picture perception (§ 4.4) - the important point being that the distal source is presented, even though apparently, within the peripersonal space of the subject (§ 4.5). Therefore, these two kinds of perceptual states – face-to-face perception and the perception of a depicted object - are not very different. This could be due to the fact that the covariation of the same geometrical properties shared by an object in perspective in a picture and a real object in perspective lead to the same representation of action properties (cfr. with § 4.4); and this is possible because the dorsal system does not distinguish between the distal cause encoded, arguably because the AIP-F5 processing exhibits the same encoding in these different situations (cfr. with § 4.4), where the location – even - in the peripersonal space of the depicted object seems to be the only relevant thing for motor related effects to be represented. So, I think I have provided enough evidence to strengthen the main claim of this chapter.

But now a question arises: if what I've said is true, what does this imply for the dorsal/ventral account of picture perception I mentioned at the beginning of my chapter? I now want to clarify the compatibility of my claims with those of Nanay and Matthen.

We saw that, concerning the commitment about (b), two positions can be considered for Nanay's account: the one in Nanay (2011a), and the one in Nanay (2015). Indeed, in his (2015) Nanay exposes his specifications for possible misunderstandings concerning (b) in relation to what he says in (2011). While in the (2011) one might understand (b) as being necessary – and this is the same position as defended by Matthen - in his (2015) he points out that (b) is not necessary in his account. My account is compatible with both positions, though the second one is the one we should prefer.

Concerning the first position from (2011), which is related to Matthen's position, my account is compatible with the dorsal/ventral account, insofar as it establishes that those depicted objects which are perceived by the subject as apparently being in his egocentric space and as apparently reachable can be represented as having action properties. This idea does not violate - but follows - the intuition that dorsal perception is involved in egocentric encoding. Nor does it conflict with the claim that we cannot really egocentrically localize a depicted object as a normal one. Indeed, there is no conflict because the dorsal representation of depicted objects that are perceived by the subject as apparently located in his egocentric space is possibly due to the experimental settings fostering the apparent peripersonal localization of these objects which are perceived as apparently reachable, as well as due to the fact that dorsal perception cannot

---

[82] Of course there are cases in which we guide action using visual consciousness. However, in most cases, dorsal perception does not attach much relative importance to contextual depth cues in situations in which action is extremely rapid or automatic (Briscoe 2009), which seems to be the usual scenario (Pacherie 2007). But sometimes, when the dorsal stream's preferred sources of spatial information are unavailable or when "there is time", dorsal perception uses the outputs from ventral processing; thus, the process might be conscious (for a complete overview see Briscoe 2009: 441). Note that, at a first step, the visuomotor system processes a gist of the motor act suitable in the given situation regardless of the fact that, in a second step, this motor act will be automatically executed or inhibited, or consciously monitored or diverted, or irrespective of whether we actually guide action consciously or not. This doesn't conflict with what I say in (§ 4.5; cfr. with 2.4.4) concerning different encodings (automatic/online) for the ventro-dorsal and the dorso-dorsal stream.

discriminate between real and depicted objects.

If we do not endorse this explanation, all the sets of evidence from vision and motor neuroscience I have reported in (§§ 4.4, 4.6) remain unexplained under the dorsal/ventral account of picture perception. Keeping in mind the dorsal lack of discrimination is useful also for the notion of egocentric localization that Nanay offers in his (2011a): egocentric localization is "the ability to localize an object in one's egocentric space" (Nanay 2011a: 468). However, as he specifies, this ability does not necessarily have to rely on the ability to interact with the object, even though this will be a good way of arguing for (b) - and thus (c). I agree that manipulation is not necessary in order to egocentrically localize an object. So, according to Nanay, in order to localize an object in one's egocentric space, we have to be able to represent the distance between the perceived object and ourselves in some way and to have expectations about how this distance can change by moving closer to, or away from, the object" (p. 469). Once again, due to its computational inability, dorsal perception represents the depicted object and its action properties on the basis of its apparent location in the peripersonal space. Of course, this does not conflict with (c) (see also § 4.4), depending on whether dorsal perception is attuned to the surface or to the depicted object[83].

For these reasons, my claim can be understood as a genuine (neuro)logical extension of the dorsal-ventral account, it being compatible with what I have said[84].

Finally, concerning the second version according to which (b) is not necessary in Nanay's account (2015), another way to read my chapter is considering it as an attempt to explain what happens when (b) is not the case.

4.7 Conclusion of the Section

I have suggested that dorsal perception can represent depicted objects as having action properties. Therefore, the reach of my proposal offers a crucial implication for the dorsal/ventral account of pictures, as well as for the debate on pictures in general. According to the accredited model, ventral perception can represent both the surface and the depicted object, but dorsal perception can represent only the surface (the picture vehicle), while it cannot represent the depicted object. But demonstrating, on my part, that also dorsal perception can represent the depicted object might entail a sort of closure of the circle[85]: both streams can represent both depicted objects (and their surfaces) and normal objects, in a very similar way[86]. This is suggested by the fact that, in the "eyes" of dorsal perception, they are almost the same distal target. But if those two streams are the total components of the visual system of humans (and other mammals), then our visual system functions in almost the same way in both picture perception and

---

[83] It has been claimed that, most of the time, when we look at pictures we ignore the surface (Levinson 1998; Lopes 1996) and that we represent both the depicted object and some of the properties of the picture surface (while we may or may not attend to the surface) (I found this discussion in Nanay 2011a: 463-464).

[84] I want to clarify that I am not arguing that it is not possible. I simply do not focus on this possibility.

[85] Crucial for this - and in line with this and with what I say in (§ 4.4) - evidence shows that action possibilities are not coded only by the dorsal stream, but by both visual pathways (Young 2006: 134; cfr. with my § 2.4).

[86] Although I have reported several differences, the point, with respect to the received view, is that also the dorsal stream can, in some way, encode depicted objects.

face-to-face perception[87]. Moreover, action properties are not properties that we can attribute only to normal objects. This doesn't mean, though, that face-to-face perception and picture perception are almost the same perceptual phenomenon; it means that they are more related than suggested by the dorsal/ventral account I am reviewing here. Moreover, a further advantage of my review is that it could be viewed as an extension of such account, while remaining within the same philosophical framework.

As argued in this chapter, both normal and depicted objects can be represented as offering action properties; thus, face-to-face perception and picture perception are not so different as concerns the visual system processing. Nevertheless, is indisputable that normal objects are perceived as present, at the visual level, in a way depicted objects cannot (Matthen 2010; Nanay 2015). According to the literature, depicted objects cannot foster the visual feeling of presence (FOP) because they cannot be represented in egocentric coordinates. The next chapter strengthens this idea, but aims to specify that egocentric localization is responsible for visual FOP because it is strictly linked to the perception of absolute depth cues that give rise to the most important aspect of visual FOP: qualitatively rich stereopsis (Vishwanath 2014) which is linked to the perception of the possibility of reliable motor interaction with the object we face with.

---

[87] The most important difference between face-to-face and picture perception is the impression of *stereopsis*, which is not possible in pictures, whose relative depth cues cannot be scaled in absolute depth cues, through which stereopsis is induced (see Vishwanath 2014). I cannot deal with this here.

5 Visual Feeling of Presence

Our everyday visual experience constantly confronts us with real things we can interact with. I look at the glass on the table, the glass shows up in my visual experience as physically present in front of me, as *really* occupying a portion of the external environment; I literally feel the "outside" presence of that "real" physical object *via* visual perceptual experience. As Matthen (2010) observes,

> "When I look down at my hands right now, it looks as if they are working on a black computer keyboard. There is something about my visual state that makes it seem as if the keyboard is really there, and that it is really black. (…) When I am relaxing in an armchair, I can close my eyes and summon up a fairly detailed and vivid image of my hands on a black keyboard. This state of visual imaging is different from my present visual state. It does not make it seem as if the keyboard is really there, nor that it ever was, or that it is really black. What is the difference between these two states? Why does the keyboard have (as I shall say) a *feeling of presence* in my present visual state, but not in the state of visual imaging?" (p. 107, 108)

I shall avoid here the issue about the relationship between visual content and the content of visual imagery (see Nanay 2014a). Rather, the *rationale* of this chapter is that the visual feeling of presence (henceforth: FOP) is a crucial feature of vision that is largely unexplored in the philosophy of perception, and poorly debated in vision neuroscience (Seth 2014, 2015). Broadly speaking, the general literature on FOP deals with three main cases. The first concerns neurological disorders regarding "those situations in which someone experiences a powerful hallucination, not explicitly sensory at all, concerning the feeling of presence of someone or something nearby, a presence felt as malevolent or benign. This feeling is about the strong and irresistible conviction that someone is there" (Sacks 2012: Chap. 15). Since I am interested in visual FOP, we can skip this case. Another case concerns the dissociations between face identification and the FOP of a familiar person despite recognizing his/her face (Dokic 2010). However, this case involves other non-visual features such emotions, familiarity and memory and thus I shall bypass this case as well. Hence, I think the best candidate[88] for an investigation about the nature of visual FOP is the remaining case, namely, the debate on picture perception, where it is widely agreed that, although depicted objects foster in the onlookeers a sort of visual experience similar to that obtained in face-to-face perception (Kulvicki 2014; Nanay 2010b, but see also 2011a; Lopes 2005; Wollheim 2003; Berger 2011), they cannot foster any FOP (Nanay 2015), or at least not as much as real objects can[89]: they show up as *not* present (Noë 2012: 84), involving a distinct presence-in-absence structure (p. 86), while face-to-face experience is comparable to touching the object (p. 87). A good and novel strategy in order to find out what makes real objects really present to us is the investigation of the nature of such difference concerning this FOP that we have in face-to-face perception and that we cannot gain from the perception of depicted objects.

The aim of this chapter is therefore to investigate the way we entertain this FOP with external objects, which we cannot entertain with depicted objects. I will suggest that FOP depends on the perception of egocentric absolute depth cues giving rise to a peculiar visual feature linked to FOP: qualitatively rich stereopsis. Since we have

---

[88] I do not mention here the case of visual imagery, or amodal completion (see Briscoe 2011; Nanay 2010a).

[89] It is one thing is to say that face-to-face perception and the perception of depicted objects do not foster the same FOP, it is another to say that picture perception cannot foster FOP at all (see § 5.6).

no account about the nature of visual FOP, my chapter constitutes a genuinely novel proposal. Indeed, to my knowledge, the only one who has talked about the difference between FOP in normal objects and that in pictures is Matthen (2005, 2010), although many others have commented on this attempt (Noë 2012; Fridland 2012; Dokic 2010; Nanay 2015). However, Matthen is interested in the semantic significance of FOP and not in the conditions under which our visual system certifies something as present/real (2010: 107-108), which is my specific interest here.

In the first part of the chapter, I explain the perceptual nature of visual FOP (§ 5.1). I suggest that FOP depends on the perception of absolute depth cues giving rise to a peculiar visual feature of FOP: qualitatively rich stereopsis (§ 5.2) - I have to point out that here I follow the notion of stereopsis developed by Vishwanath and colleagues (see Vishwanath 2014), who offered the most recent theory of stereopsis in the literature of vision science; I show that FOP about the picture vehicle/surface hinders the FOP about the depicted object (§ 5.3). I also suggest that egocentric depth is linked to the observer's knowledge of the depth relations scaled in a meaningful way to the observer's motor action space (§ 5.4). I report the links between the neurophysiological/psychophysics level and the phenomenological level concerning visual FOP (§ 5.5) and offer some important remarks about the experimental evidence I use for my theory (§ 5.6). Finally, in (§ 5.7) I frame the debate on visual FOP in the empirical framework of the TVSM, insofar as the debate on vision generally follows this model. Thus, I address the neural correlates of FOP with respect to the TVSM.

5.1 FOP and its Perceptual Cues

I start here from the widely agreed empirical evidence that, while we can get an unambiguous perception of depth and 3-D structure during monocular perception or during the visual perception of a pictorial image of a 3-D scene (Koenderink and van Doorn 2003; Koenderink, van Doorn and Wagemans 2011; Vishwanath, Girshick and Banks 2005; quoted in Vishwanath 2014: 154; Vishwanath 2010, 2014; Albertazzi et al. 2010), face-to-face binocular perception of depth is qualitatively different from any other visual situation. This difference is due to the vivid impression of tangible solid form and immersive negative space, namely, stereoscopic vision, or stereopsis (Vishwanath 2014: 151, 154, 156, 159, 171). From this definition, it is easy to understand that stereopsis deals with the visual features we engage with when we experience visual FOP: presence of stereopsis means presence of FOP[90]. In other words, if we want to look for a description of FOP, we need to investigate the nature of the visual attribute that leads our visual system to "establish" that a thing is real. Stereopsis is the best candidate for defining this visual attribute. Accordingly, we have to consider our most recent theory about stereopsis, which is able to collect all recent data from vision neuroscience. Here, in order to explain the nature of FOP, I follow the most recent account on stereopsis, which is proposed by Vishwanath and colleagues (see Vishwanath 2014).

5.2 Stereoscopic Vision

Following to the most recent theory of stereopsis (Vishwanath 2014), which is based on the *Absolute Depth Scaling Hypothesis*, the perception of 3-D structure depends only on

---

[90] The fact that we perceive a "real" external reality due to stereopsis is uncontroversial in the neuroscientific literature (Kandel et al. 2013). However, defining the visual process through which we reach stereopsis is another issue (see my § 5.6) and using this account to explain FOP is something neuroscience seems not to care about.

relative depth estimates, while the impression of stereopsis depends on the perception of egocentric absolute depth values (p. 155). *Relative depth* concerns the fact that an observer understands the metric relative to depth relations within and among objects (3D surface shape, surface slant, depth order, etc.) and the ratios of distances or depths among points in the visual field - in the example of Vishwanath (2010: 222). E*gocentric absolute depth* concerns the fact that the observer understands the depth relations scaled in some meaningful way to his/her own actions (egocentrically scaled depth separation). Stereopsis is primarily based on the relation between the vivid impression of solid object shape and the impression of a palpable invisible negative space between objects, both responsible for the impression of "real separation" between points in depth, either within an object (perception of solid shape) or between objects (the impression of a palpable invisible space). This impression of 'real separation in depth' varies as a function of the precision of absolute depth estimates.

Crucially, relative depth (3D surface shape, slant, depth order) and absolute (egocentric) depth are separate, dissociable perceptual constructs, from both a phenomenal and a psychophysical standpoint, insofar as they differ concerning the cue-combination. Indeed, depth cues only provide relative depth information that has to be scaled by egocentric distance information to derive absolute depth (Vishwanath 2014: 154). What is important here is that this visual scaling is not possible in picture perception, insofar as pictures permit us to perceive only relative depth cues. Indeed, since we can perceive metric relative depth structure (3-D object shape and layout) in conditions that fail to induce stereopsis (Vishwanath 2014: 171), we can thus perceive relative depth when looking at depicted objects, even though, without distance information, these cues cannot be scaled to derive absolute depth or size (Vishwanath 2014: 158-159): distance signals such as vergence, accommodation and so on specify distance from the picture surface and not from its pictorial contents (p. 224 for a review).

To sum up, pictures and real scenes differ in terms of the presence of stereopsis, insofar as the impression of real separation in depth varies as a function of the precision of absolute depth estimates and the value of this function is null when we perceive depicted objects (p. 170-171). During binocular picture perception no combination of retinal or extra-retinal information can specify the absolute depth separation between objects in the pictorial scene (Vishwanath 2014: 158) and thus we lack the sense of an immersive negative space between them. More importantly, there is no optical distance cue specifying the distance of pictorial objects from the onlooker, who lacks the egocentric, viewer-relative locations of its objects (Vishwanath 2014: 158; Vishwanath 2010: 202-206) and the vivid impression of tangible solid form, which is possible in ordinary seeing (Vishwanath 2010: 223-224; Vishwanath and Hibbard 2013). Thus, we lack FOP

This idea that FOP depends on the perception of egocentric absolute depth is perfectly compatible with Matthen's account (see 2005: 315-316; 2010).

5.3 Surfaces

Another crucial reason why FOP is not experienced with pictures is that surface invisibility plays a role in the induction of stereopsis (Vishwanath 2014: 164): FOP about the picture vehicle/surface hinders the FOP about the depicted object (Vishwanath 2014: 158-175): "in the absence of visible picture surfaces, the brain attributes the accommodation response to the pictorial objects, and assigns any associated distance information to them, allowing absolute depth values to be derived: viewing conditions yielding only a partial reduction in surface visibility results in a greater variability in the assignment of distance information to the pictorial objects and a correspondingly reduced sense of stereopsis" (Vishwanath 2014: 159-160). So, the relation between surface perception and the perception of the depicted object is crucial: when the picture surface is not visible, distance information is assigned to the pictorial object allowing a derivation of size and absolute depth values (I do not want to venture into technical details, see Vishwanath 2014: 160). In other words, the fact that we perceive the pictorial surface, leads to the lack of the impression of FOP during picture perception. Since, unlike experimental conditions with different settings built *ad hoc* for the experiments (see § 5.6), during ordinary seeing we always perceptually represent the picture surface, then we never run into FOP with pictures. This seems to be a form of inflection, that is to say the phenomenon according to which some of the visual features of the picture's vehicle are relevant in characterizing features of the picture's subject (as seen in that vehicle) (Voltolini 2013; Nanay 2005). This deals with the twofoldness of pictorial perception (seeing-in), namely, the fact that our visual system perceptually represents both the surface/vehicle of the picture and the depicted object (Wollheim 1998; Nanay 2005, 2011), something uncontroversial in the debate on pictures (for an analysis of the nature of the visual representations involved in seeing-in see Nanay 2011: 462-464; Lopes 2005: Chap. 1) - I'll be back to this point in (§ 5.7).

5.4 Presence in Motor Space

An important point linked with what I have said above is that we perceive the objects in the external environment as affording motor interaction, an impression that pictorial objects cannot foster (but see § 5.7). This is because absolute depth perception linked to stereopsis concerns the observer's perception of the depth relations and spatial parameter scaled in a meaningful way for potential motor interaction, while relative depth perception only relates to the capacity to visually match the 3-D shapes of objects (Vishwanath 2014: 174; Vishwanath 2010: 232). There is, indeed, significant evidence that stereopsis is important for visually-guided grasping (Watt and Bradshaw 2003) – this is in line with Matthen's idea (2005, 2010) of a link between FOP, egocentric localization and the perception of action possibilities.

5.5 Phenomenological Reports

We have also experimental results showing that the phenomenological reports concerning the impression of stereopsis in different experimental settings are in line with the results I have reported in the previous sections (cfr. with § 5.6). Naïve and expert observers alike suggest that stereopsis is associated with several specific qualitative characteristics. First, a sense of real separation in depth between points on an object or between objects themselves; the former yields the visual impression of 'solidity', the latter the impression of a 'real' negative space; second, a sense of clarity and visual sharpness, which does not depend on depth, and which leads to a more enhanced impression of color (variation) and material qualities (glossiness, shininess, roughness etc.); third, a sense of tangibility (impression of the manipulability of a material object) and of spatial immersion (impression of the capacity to move through a palpable negative space) (Vishwanath 2014: 174; Barry 2009; Vishwanath and Hibbard 2013). These three features realize a vivid sense of protrusion in which a tangible solid object reaches or looms out through the negative space toward the observer. These features give us the sense of 'reality' linked to binocular stereopsis, often called the "plastic effect" (see Vishwanath 2010: 224, 225; Vishwanath and Hibbard 2010), related to a sense of capacity to interact with real objects in a 3-D space; this is precisely the effect FOP gives us.

An important thing: we saw that FOP is massively linked to motor space. It's not by chance that the strongest plastic effect is found within the (peri)*personal action space*, while it diminishes going beyond this space into distant space (*vista space*), which fosters little or no perceptible plastic effect: observers comment on how distant landscapes often appear "pictorial" (Vishwanath 2010: 228, 225) - this has been claimed by Matthen as well (2005: 322). This is because the reliability of sources of egocentric distance information and absolute depth estimation relying on distance estimates (vergence, accommodation, disparity) reduces when distance increases (Vishwanath 2010: 225). For the same reason, we have a complete absence of the plastic effect in the binocular viewing of pictures, in which, as said, absolute scaling is impossible (Vishwanath 2010: 227). I think that, beside the neuroscientific accounts, everybody - I refer to healthy individuals - realizes that a depicted glass of water cannot induce in our visual experience the same FOP that a real glass solicits in us.

These results suggest that there are so many differences, between face-to-face seeing and picture seeing, at a perceptual level, concerning those necessary perceptual cues FOP relies on (stereopsis, egocentric depth estimates and the plastic effect), that we do not need to invoke further mental states in order to establish their difference in terms of FOP. An important implication of this view is that, differently from the widely agreed idea that feelings are metacognitive or cognitive states to be attached to a perceptual state (Matthen 2010, Martin and Dokic 2013; Dokic 2012), FOP is just the result of a perceptual state: real-scene vision is actuality-committing (this expression is from Matthen 2010: 3), in a way that picture seeing is not insofar as it has access to this perceptual stereoscopic information that only normal objects can provide our visual system with and that we cannot gain from depicted objects. Thus, I continue to talk about a "feeling" and not about a "perception" of presence, just for the sake of continuity with the philosophical literature. This feeling is not something to attach to a perceptual state.

5.6 An Important Remark about the Experimental Results concerning Stereopsis

For coherence with the results in vision science provided by Vishwanath account, I shall now specify important implications about this theory of stereopsis I have mentioned. It is widely agreed that, in binocular visual perception, different views of the object are simultaneously seen with each eye (binocular parallax) and, during self-motion, different views of the object or scene are seen in temporal sequence (motion parallax): simultaneous or successive presentation of different perspectives of an object (visual parallax) are usually responsible for stereopsis (I rephrase Vishwanath 2014: 152). Against the common view that stereopsis emerges from binocular disparity, or at least from motion parallax, this theory claims that monocular stereopsis - and the related plastic effect – is possible and can be induced by reducing the factors responsible for the visibility (perceptual representation), of the picture surface (Vishwanath 2014: 158) (cfr. with my § 5.3), for example, when absolute depth scaling is made possible in picture while monocularly viewing a picture through an aperture[91] (Vishwanath 2014: 160; for a review see Vishwanath and Hibbard 2013). This implies that, contrarily to what is believed, stereopsis is not a by-product of binocular vision (Ponce and Born 2008) and that we should distinguish between binocular depth perception (the capacity to perceive

---

[91] Blur gradients can induce stereopsis in different not experimental situations that go far beyond ordinary seeing, such as with monocular viewing or even if when surface visibility is present (Vishwanath 2010: 227).

quantitative depth relations using the visual information from two eyes) and stereopsis (qualitative vividness of depth that is often obtained as a result of this capacity) (Vishwanath and Hibbard 2013). Indeed, "stereopsis from binocular image pairs with disparity should be significantly stronger than monocular stereopsis (viewing through an aperture), not because disparity is the main cue for the impression of stereopsis but because disparity can be scaled by binocular convergence, which is a more reliable distance cue than accommodation (Howard & Rogers 2002), and should yield a more precise estimate of absolute depth" (Vishwanath and Hibbard 2013: 1684). This implies that the distinction concerning stereopsis is not between pictorial and real space, but between relative and egocentrically scaled depth, insofar as the latter can be reliably present or absent in pictures and real scenes depending on the available egocentric distance information (I have rephrased Vishwanath 2010: 230; see also p. 226; see also Vishwanath 2007, 2008; Mausfeld 2003). Indeed, the distinction between relative and absolute depth perception depends on the differences in depth perception between pictures and real objects, monocular and binocular viewing, and vista space and peripersonal-action space (Vishwanath 2010, 2014). This undermines the idea that, in absolute, either we entertain with FOP, or we do not and there are no middle states. Indeed, depth relief of stereopsis decreases in the following order: binocular viewing of a real or stereoscopic images of the object; monocular viewing of a picture or monocular viewing of a real object; binocular viewing of a picture. This is the rank of the reports about reduced impression of stereopsis given by participants in the experimental settings (see Vishwanath 2014: 173). The perception of relative depth (3D surface shape, slant, layout) during picture seeing (monocularly or binocularly) is *more or less veridical* although it may not be *metrical in the absolute sense*" (Vishwanath 2010: 223). Veridical means that the perception of the relative depth in a picture is *more or less* the same we see during binocular vision of real objects (p. 223). The same holds for monocular vision of a real object (p. 224). From this evidence it follows that, if FOP depends on (the degrees of) stereopsis, then FOP can come in degrees as well, with respect to the visual situations we are dealing with in different experimental settings - an analysis of ordinary seeing follows what I have said all along (§ 5.1). The reason why I do not need to take into account this fact is that stereopsis is possible in pictures due to particular experimental settings we cannot be in in everyday life. But my investigation here is about the difference in stereopsis and FOP concerning two ordinary cases[92]: binocular face-to-face seeing and binocular picture seeing. Therefore, the fact that the real nature of stereopsis does not depend on binocular seeing, insofar as it can be fostered in particular situations concerning monocular viewing of pictures we never deal with in ordinary life scenarios is not relevant for the purpose of this chapter. It would be interesting to investigate FOP taking into account the real nature of stereopsis, through different experimental settings, following the philosophical implications of this new theory of stereopsis. But this is not the place to do so.

To sum up, an important consequence of this evidence is that FOP is not necessarily attributed to real objects only. Indeed, the explanation I endorse - that our visual system "defines" an object as being really there only when particular cues are perceivable – is able to explain why it is possible for us to perceive the cues necessary for FOP also during picture seeing, but in particular experimental settings, which are far from our ordinary situations in everyday life. So, my explanation has the advantage of

---

[92] This focus on ordinary seeing is unproblematic. Even Matthen's claim is about veridical perceptual situations and is that normal scene vision is actuality-committing while pictorial vision is not (2010: 15, 19); thus, Matthen's idea is about ordinary, everyday life visual states (2010: 14), without considering illusions, imagery, dreaming, hallucinations, etc. - the same holds for Martin and Dokic (2013).

establishing the difference between picture and face-to-face ordinary (binocular) seeing, while remaining also compatible with this evidence that goes beyond ordinary seeing and considers the possibility of FOP for pictures. However, as said, here I am interested in the difference, concerning visual FOP, between binocular picture seeing and binocular face-to-face seeing.

5.7 FOP and the TVSM

Now, since the debate on vision – as well as the debate on picture perception (see Nanay 2011a, 2015) - is framed in the TVSM, the relation between visual FOP and the TVSM needs to be addressed.

As seen in chapter 2 and chapter 4, the TVSM suggests a twofold anatomo-functional dissociation of the visual pathways (Milner and Goodale 1995/2006): a ventral stream responsible for identification and recognition, allowing us to perceive from an allocentric frame of reference, and a dorsal stream is involved in guiding action, allowing us to perceive from an egocentric frame of reference (Jacob and Jeannerod 2003). Usually, ventral perception is taken to constitute the mark of conscious vision, while dorsal perception remains completely subpersonal. Indeed, the streams can be dissociated due to cortical lesions. Lesions in the dorsal stream (the occipito-parietal network from the primary visual cortex to the posterior parietal cortex) impair one's ability to use what one sees to guide action (optic ataxia), but not object recognition; lesions in the ventral stream (the occipito-temporal network from the primary visual cortex to the infertemporal cortex) impair one's ability to recognize things in the visual world (visual agnosia), but not action guidance (Jacob and Jeannerod 2003). Moreover, we have behavioral studies of normal subjects involving visual illusions that can deceive the ventral stream but not the dorsal one; thus, it seems that, contrarily to ventral perception, dorsal perception is completely impenetrable by consciousness (Bruno and Battaglini 2008; McIntosh and Schenk 2009).

Here I place my account on visual FOP in relation to the framework given by the TVSM in vision neuroscience by explaining the relation between FOP and the neural coordinates the TVSM relies on. Of course, I do not take my explanation to be complete for neuroscience: it only aims to give a sketch of what is the crucial relation between FOP and the two streams, which turns out to be very important for a philosophical theory of visual FOP that is based on empirical results. Of course, I do not take my explanation to be complete for neuroscience: it only aims to give a sketch of what probably is the relation between FOP and the two streams. I argue that FOP depends on interstream interaction: ventral perception is crucial for high perceptual reconstruction, while dorsal perception is crucial for visuomotor perception; two processes - as we saw in this chapter - crucial for FOP.

First of all, as explained in the previous chapters, we know that dorsal perception is the one crucially involved in constructing visuomotor representations for suitable motor interaction. However, we also know, as shown in (§§ 2.5, 4.6), that dorsal perception is sensitive, to some extent, to depth cues apparently exhibited by depicted objects (Gonzalez et al. 2008; Berryhill and Olson 2009; Marotta and Goodale 2001; Taira et al. 2000; Tsutsui et al. 2002, 2005; James et al. 2002, Sakata et al. 2003; Nelissen et al. 2009; Chinellato and del Pobil 2015) and seems to discriminate between images of depicted graspable and non-graspable objects (Rice et al. 2007). However, as we saw in (§ 5.1), all those 3-D depth cues are not sufficient for stereopsis and visual FOP. So, even if dorsal perception is the one involved in egocentric localization (Milner and Goodale

2006), it cannot be, alone, responsible for FOP. Consider also that dorsal perception cannot discern between real graspable objects and graspable objects in pictures, insofar as dorsal perception cannot construct a *complete* 3D structural description of the target object (Westwood *et al.* 2002), which is, nonetheless, necessary for response *selection*, in order to detect the action afforded by an object, or in the case of pictures, to understand that there is no possible interaction[93]. Accordingly, ventral perception plays a key role in object recognition (Turnbull, Driver and Mcarthy 2004; Kravitz et al. 2013) and response selection, distinguishing between 3D objects and 2D images of objects by detecting conflict between various visual cues[94] and selecting different visuomotor strategies encoded by dorsal perception, concerning whether we are dealing with a 2D image or a 3D object. That is, ventral perception is the one able to perceive qualitative visual features in order to construct complete volumetric representations of objects (for a review see Westwood et al. 2002)[95] – and this is really important, if we do not want to attempt to reach for non-graspable objects, risking to hit a wall – (see chapter 4). All this is a starting point to suggest an interstream interplay in order to obtain FOP: dorsal perception for the computation of motor action in an egocentric coordinates and ventral perception for high resolution volumetric reconstruction and response selection.

Accordingly, from chapter 4 we know that the same dorsal representations active when we look at manipulable objects are active for depicted objects as well. Recall that, indeed, several portions of the ventro-dorsal stream, which is the crucial cortical portion of the dorsal stream involved in the transformation of intrinsic object properties into action properties with the consequent translation of those properties in suitable motor acts, are activated during picture perception: F5 in the most rostral part of the ventral premotor cortex (Chao and Martin 2000; see also Buccino et al. 2009; Costantini et al. 2010; Proverbio et al. 2011; Grezes and Decety 2002; for a review see Zipoli Caiani 2013), parietal neurons in the dorsal stream and the anterior intraparietal area (AIP) (Romero et al. 2014; Pani et al. 2014). Therefore, the ventral premotor cortex, which is an end-stage projection of the dorsal stream, is highly selective for manipulable objects observation with respect to non-manipulable objects (Gerlach et al. 2002; Kellenbach et al. 2003). As explained in (§ 4.5), this is likely to happen when the depicted object is apparently presented in the peripersonal space of the observer and is apparently reachable.

Of course, this does not mean that we consciously perceive depicted objects as affording a real possibility of action, but only that dorsal perception responds to those similar geometrical patterns exhibited by depicted objects, that usually instantiate the arrangement of an action property in a normal object – (for a review see §§ 4.4, 4.5, 4.6)[96]. This evidence that dorsal perception responds to depicted objects undermines the

---

[93] Volumetric object representation is necessary for the visual control of grip formation and response selection, to ensure that we do not attempt to reach for objects that cannot be grasped.

[94] It is computationally efficient for one visual system to handle both response selection and object recognition: both require complete/detailed information about 3D object structure (Goodale and Milner 1992).

[95] For the difference between shape perception and volumetric object recognition, which is crucial here in the light of what I said in (§ 5.1) see (Briscoe 2008).

[96] Note that, in line with what I have said in (§ 4.5), dorsal response to depicted objects is active only when the surface of the depicted object is presented within the peripersonal space of the observer (Costantini et al. 2010; Cardellicchio et al. 2011). This is arguably because, as said, dorsal perception cannot distinguish between normal and depicted objects, and then responds to those objects apparently presented within the peripersonal space of the observer, regardless of the real nature of the distal target - (for a review see §§ 4.4). This confirms the idea – expressed in (§§ 2.3.3, 4.5) - that the possibility of interaction and

possibility that FOP depends on dorsal perception only: if dorsal perception is involved in the representation of both depicted and normal objects, and if FOP is experienced only in the case of (binocular) perception of normal objects, but not in the case of (binocular) perception of normal pictures, dorsal perception cannot be the particular perceptual state responsible for the visual FOP. All this is in accordance with what I have said in (§ 5.4). Indeed, when it comes to conscious visual perception of normal objects, we need both dorsal perception constructing visuomotor representations for interaction and ventral perception selecting the real object in the environment by performing the response selection. In accordance with all what I said, we know that the conscious egocentric localization linked to the conscious perception of possibility of motor interaction – crucial for FOP - is due to interstream interaction (Briscoe 2009: 6.1–6.4), insofar as dorsal perception can have ready access to sources of spatial content in the ventral stream and makes use of contextual depth cues and other sources of 3-D spatial information in the ventral stream – for a complete overview about the neural processing concerning the interaction between the visual representations involved in stereoscopic perception and the MRs allowing us to interact with the environment see (Chinellato and del Pobil 2015, Romero et al. 2013, Theys et al. 2015).

Finally, the dependence of FOP on both streams is suggested by the fact that dorsal lesions produce deficits in the conscious awareness of the quality of objects presented within peripersonal space (Gallese 2007) and that the ventro-dorsal stream, (related to the inferior parietal lobule), which is the most ventral portion of the dorsal stream, engaging in several interconnections with the ventral stream (see Fogassi and Luppino 2005), is crucial for visual awareness and information specified in egocentric coordinates that requires higher ventral processing (Gallese 2007; Jacob and Jeannerod 2003: 252–255; Brogaard 2011: 1094). Also, while dorsal processing alone is not responsible for visual consciousness, given all the visual cues it cannot have access to, ventral processing is not sufficient to obtain perception without parietal spatial processing, nor can it encode motor interaction without dorsal processing (Rizzolatti and Matelli 2003) - this explains the interplay of both streams and the related interplay, expressed above, of the response selection and of the visuomotor encoding - for a review of these experimental results see above (§ 2.4.5). In turn, this suggests that FOP is a result of the interaction between the streams.

Of course, from what I said in chapter 2, especially (§ 2.4), my point about interstream interactions is not problematic, insofar as it has been shown, going beyond the TVSM, that any sophisticated visual behavior – *a fortiori*, two complex visual processes like stereopsis and FOP - requires collaboration between both streams at every stage of the process (Schenk and McIntosh, 2010; Kravitz et al. 2011, 2013; Deco et al. 2004; Grill-Spector et al. 2001; Gallese et al. 1999; Gallese 2007: 1). Accordingly, recall that several chunks of the dorsal stream are interconnected with inferotemporal areas (TEm, TE and TEO) of the ventral stream and the visuomotor transformations of objects properties in motor commands rely on the analysis of both physical properties (pragmatic analysis) and object identity (semantic analysis) (see Fogassi and Luppino 2005, Rizzolatti and Matelli 2003 and Gallese 2007; Rizzolatti, Sinigaglia 2008: 36-38; Cisek 2007; Kandel et al. 2013: Chap. 19) – for a review see (§§ 2.4.1, 2.4.3).

Summing up, several empirical arguments exclude that FOP depends on the activity of a singular stream, insofar as a lot of cues that FOP depends on cannot be

---

peripersonal encoding march in step (see Turella and Lignau 2014), but, an important point, in the case of dorsal perception only, they are due to subpersonal processing - (for a review see §§ 4.5, 4.6).

processed only by a singular stream. On the one hand, ventral perception, with its response selection and volumetric reconstruction, is crucial for FOP, which depends on the high perceptual resolution linked to stereopsis and absolute depth cues – and, indeed, the ventral stream goes from the primary visual cortex V1 until the inferotemporal cortex, a crucial pathway where the highest informational processing about qualitative rich visual cues are computed; on the other hand, dorsal perception, is crucial in encoding visually guided action performance in egocentric coordinates. In line with a recent proposal by Briscoe (2009), it is interstream interaction that allows visual experience to represent 3-D space around the perceiver using an egocentric frame of reference and in order to gain a constitutive link between the spatial contents of visual experience and the perceiver's bodily actions.

A particular point for a theory of FOP needs to be considered here. Sometimes we have the chance of being deceived by trompe l'oeil pictures for a moment, and thus we experience FOP.

Some authors do not consider trompe l'oeil pictures as a genuine case of pictures (Wollheim 1987, 1998; for a discussion see Nanay 2015, Voltolini 2013; Hecht et al. 2003), and even those who underlined the impressive sense of reality fostered by them (and even by other non trompe l'oeil paintings) argued that pictures foster "blind" visuomotor experiences, corresponding not to actual movements, but to "virtual" movements internal to acts of perception, even though the subject perceives a full sense of the physicality of things (Smith 2014: 102). However, while some claimed (Wollheim 1987, 1998) that trompe l'oeil paintings are not pictures because they do not fall under the visual phenomenon of seeing-in, or two-foldness, which is the perceptual mark of figurativity - what makes a representation pictorial (see Voltolini 2013: Sec. 2; Lopes 2005 pp. 34-36; see also Nanay 2005, 2010b) - it has been suggested that this argument might be undermined (see Feagin 1998: Sec. 2; Levinson 1998; see also Nanay 2011a: footnote 11). Stepping back from this debate, what is important for the point I am discussing here is that it is widely agreed that, even though for just a moment, we are in fact fooled by trompe l'oeil paintings (see Lopes 2005, in particular Sec. 1). So, I endorse that trompe l'oeil pictures foster in us the FOP for a while, until we are not able to recognize that they are pictures and not real scenes insofar as our visual system is temporally fooled. For this reason, I have to account for this special kind of picture rather than avoiding this point by arguing that tromp l'oeil pictures are not actually pictures. Accordingly, the FOP experienced during the perception of trompe l'oeil pictures can be taken into account in my theory.

First, I already said that particular experimental settings given by particular pictures are able to foster FOP; this leads me to consider the case of trompe l'oeil as one of the particular cases reported above concerning particular pictures. Furthermore, I said that experimental evidence does not show that pictures - even normal pictures - cannot induce stereopsis and FOP, but just that, in ordinary visual experience, that is, without particular kinds of pictures and/or experimental settings, real objects can give us access to absolute depth cues that pictures cannot offer. This is not in conflict with the fact that trompe l'oeil pictures, in their quality as special kinds of pictures, are able to foster in us, even if just for a while, the impression of the possibility of absolute depth scaling of the depicted object, which is a crucial feature in order to obtain stereopsis: this is in line with the idea by Nanay that, when we look at trompe l'oeil pictures we do (mistakenly) represent the depicted object as having a spatial location in our egocentric space (2015: 194; see also Lopes 2005), or, in the expression I used above, we have the impression of an absolute depth scaling. This depends on the fact that we have, for a moment, no recognition of the picture surface (see my § 5.3), as commonly suggested in the philosophical literature (Lopes 2005: 35-36; see also Voltolini 2013; a good review is offered by Hecht et al. 2003), presumably because the chunks of our visual system involved in high recognition - which are mostly, but not exclusively, linked to ventral vision, insofar as ventral vision is, as said, highly interconnected with the dorsal one - are not attuned to the surface in this perceptual situation (see the related empirical point in my § 5.3), in accordance with the argument given by Nanay (2015: 194). All this is in line with my view that, with respect to the TVSM, FOP depends on interstream interactions, insofar as when entertained with FOP - and with the perception of the possibility of reliable motor interaction (which, in the case of pictures, is only apparent) - we dorsally as well as ventrally represent the object, even in the case of a depicted object as a trompe l'oeil picture (Nanay 2015: 194) - however, since dorsal perception represents even normal (that is, non trompe l'oeil) depicted objects, what I take to be peculiar of the occurrence of FOP in the case of trompe l'oeil pictures is the missing representation of the surface; I would like to be clear on the fact that we should avoid the idea that the difference concerning the occurrence of FOP between binocular viewing of trompe l'oeil pictures and binocular viewing of normal pictures is given by the fact that – in addition to the lack of the recognition of the surface - the dorsal component (the visuomotor system) comes into play in the case of trompe l'oeil, while it was not active with normal depicted objects; indeed, as said – and this is a crucial point here – the dorsal visual processing is always responsive to depicted objects, because it cannot distinguish between a normal and a depicted object; therefore, maintaining that the dorsal response is important for FOP, my idea is that FOP depends on the fact, I mentioned above, that we (the processing in our visual brain subserving high level visual object recognition, see above) cannot properly represent the picture surface (see my § 5.3).

All I said is perfectly compatible with the idea that the perception of reliable motor interaction is not only due to dorsal vision, but to the dorsal/ventral interplay: if ventral vision is not representing the surface, it is likely that the response selection, which is due to ventral vision, is triggered, because the trompe l'oeil depicted object is not perceived as such and, as said, the subject (her/his visual system) "has the impression" of the possibility of absolute depth scaling of the depicted object, which is a crucial feature in order to obtain stereopsis; thus, while, on the one hand, dorsal vision cannot distinguish between normal and depicted objects, on the other hand, in the case of trompe l'oeil depicted objects, even ventral vision cannot "realize" - even though for just a moment - that we are dealing with a picture and, thus, it triggers response selection. Note that, as said in (§§ 2.5, 4.6), the response selection is linked to the motor act we want perform, which is stored in the quiver of the visuomotor system due to the dorsal response, even before that ventral vision has computed whether the object is "real" or not. Indeed, following what I said in (§§ 2.5, 4.6), before ventral volumetric reconstruction a motor act computed on the basis of the 2-D geometrical properties of the target is already stored in our motor quiver, regardless of the fact that actual overt interaction will follow. However, in the case of trompe l'oeil depicted objects, the perception of the possibility of motor performance, which is due to the complex dorsal/ventral interplay I described here, is the result of the momentary illusory impression of the possibility of absolute depth scaling of the depicted object. But, still, we have to bear in mind that, even normal pictures can foster FOP in particular experimental settings. Indeed, my general claim is that the impression of FOP can be fostered by particular pictures, as well as normal pictures in particular experimental settings, but that, in ordinary visual scenarios, that is, in ordinary binocular seeing of normal (non trompe l'oeil) pictures – even when the picture is in the peripersonal space of the onlooker – and in binocular seeing of normal objects, the difference between (usual, non trompe l'oeil) picture perception and face-to-face perception is the possibility of stereoscopic perception of absolute egocentric depth scaling, which is possible with normal objects in a way it is not with pictures. This seems to be in line with the insights of the philosophical accounts of picture perception about tromp l'oeil pictures (Lopes 2005; see also Voltolini 2013; Feagin 1998; Levinson 1998) as well as with the analysis of the behavior of our visual system(s) during picture perception (Nanay 2011a), coming from our best attempt to reconcile the neuroscientific knowledge about pictures with that of the philosophical literature (Nanay 2015; I cannot review all the literature here, but see Rogers 2003; Niederée and Heyer 2003; Koenderink and van Doorn 2003; Mausfeld 2003; Cutting 2003; Hecht et al. 2003).

Finally, it is worth noting that my account is not in conflict with that of Matthen. Matthen holds that egocentric localization - and the related perception of reliable motor interaction - is responsible for FOP, and this is in line with what I said in (§ 5.1). Also, for Matthen it is motion-guiding vision that is responsible for FOP, egocentric localization and the perception of the possibility of motor interaction. For him motion-guiding vision denotes the ensemble of the functional visual processes allowing motor interaction. The fact that for him FOP depends on motion-guiding vision is not a problem here, if we assume that motion-guiding vision is the result of interstream functional interaction. And this is possible because he talks about, respectively, descriptive vision and motion-guiding vision and he clearly specifies that he is not concerned with the anatomical details of these data-streams, but only with a broad functional difference and the (double) dissociability of these functions (2010: Sec. 5).

5.8 Conclusion of the Section

In conclusion, visual FOP depends on the perception of absolute depth cues giving rise to a peculiar visual feature of FOP: qualitatively rich stereopsis. I followed the notion of stereopsis developed by Vishwanath and colleagues (see Vishwanath 2014), who offered the most recent theory of stereopsis in the literature of vision science. Also, the visual FOP about the picture vehicle/surface hinders the FOP about the depicted object. Crucially, egocentric depth is linked to the observer's knowledge of the depth relations scaled in a meaningful way to the observer's motor action space. All I said is confirmed by links between the neurophysiological/psychophysics level and the phenomenological level concerning visual FOP. Finally, I suggested that FOP depends on the interplay between the ventral visual stream and the dorsal visual stream. Accordingly, my account is neither in conflict with the idea that dorsal perception is crucial for visual FOP (Nanay 2015), nor with the idea that dorsal perception is not sufficient for FOP (Noë 2012). Indeed, not only I stressed that dorsal perception is important in order to perceptually obtain FOP, but also it is the interplay with ventral perception that gives rise to the peculiar visual FOP we entertain with normal objects.

References

- Adamo, M., Ferber S., (2009), A picture says more than a thousand words: Behavioural and ERP evidence for attentional enhancements due to action affordances, *Neuropsychologia*, doi: 10.1016/j.neuropsychologia.2008.07.009.
- Aglioti, S., DeSouza J. F. X., and Goodale M. A., (1995), Size-contrast illusions deceive the eye but not the hand, *Current Biology*, 5, pp. 679–85.
- Albertazzi, L., van Tonder G. J., and Vishwanath D., (2010), *Perception beyond Inference. The Information Content of Visual Processes,* The MIT Press.
- Algom, D., Chajut E., and Lev S., (2004), A rational look at the emotional stroop phenomenon: a generic slowdown, not a stroop effect, *Journal of Experimental Psychology, Gen.* 133, pp. 323–338, doi: 10.1037/0096-3445.133.3.323.
- Ambrosini, E., Costantini M., (2013), Handles lost in non-reachable space, *Experimental Brain Research*, 229 (2), pp. 197-202.
- Andersen, R. A., Andersen K. N., Hwang E. J., and Hauschild M., (2014), Optic ataxia: from Balint's syndrome to the parietal reach region, *Neuron* 81, 967–983, doi: 10.1016/j.neuron.2014.02.025.
- Anelli, F., Borghi A. M., Nicoletti R., (2012), Grasping the pain: Motor resonance with dangerous affordances, *Consciousness and Cognition*, 21, pp. 1627–1639.
- Anelli, F., Nicoletti R., Bolzani R., and Borghi A. M., (2013a), Keep away from danger: Dangerous objects in dynamic and static situations, *Frontiers in Human Neuroscience,* 7:344, doi: 10.3389/fnhum.2013.00344.
- Anelli, F., Ranzini M., Nicoletti R., and Borghi A. M., (2013b), Perceiving object dangerousness: An escape from pain? *Experimental Brain Research,* 228, pp. 457–466, doi: 10. 1007/s00221-013-3577-2.
- Anscombe, G. E. M. (2000), *Intention*, Cambridge, MA: Harvard University Press.
- Arnold, M., (1960), *Emotion and Personality*, New York: Columbia University Press.
- Bach, K., (1978), A Representational Theory of Action, *Philosophical Studies*, 34.
- Barlassina, L., Newen A., (2013), The Role of Bodily Perception in Emotion: In Defense of an Impure Somatic Theory, *Philosophy and Phenomenological Research*, 89 (3), pp. 637–678, doi: 10.1111/phpr.12041.
- Barrett, L. F. and Bar L. F., (2009), See it with feeling: affective predictions during object perception, *Philosophical Transactions of the Royal Society*, 364, pp. 1325–1334 doi: 10.1098/rstb.2008.0312.
- Barry, S., (2009), *Fixing my gaze*, New York, NY: Basic Books.
- Baumann, M. A., Fluet M. C., Scherberger H., (2009), Context-specific grasp movement representation in the macaque anterior intraparietal area, *Journal of Neuroscience*, 29, pp. 6436–6448.
- Bennett M., and Hacker P. M. S., (2003), *Philosophical Foundations of Neuroscience,* Wiley-Blackwell, Oxford.
- Bennett, M. Dennett D. Hacker P. M. S. Searle J., (2007), *Neuroscience and Philosophy. Brain, Mind, and Language*, Columbia University Press, NewYork.
- Bennett, M., and Hacker P. M. S., (2012), *History of Cognitive Neuroscience*, Wiley-Blackwell, Oxford.
- Berger, J., (2011), *The Sense of Sight*, New York: Pantheon.
- Berryhill, M. E., and Olson I. R., (2009), The representation of object distance: Evidence from neuroimaging and neuropsychology, *Frontiers in Human Neuroscience*, 3, pp. 1-9.

- Bickle, J., (2013), *The Oxford Handbook of Philosophy and Neuroscience*, Oxford University Press, Oxford.
- Bishop, J. M., Martin A. O., (Eds.) (2014), *Contemporary Sensorimotor Theory, Studies in Applied Philosophy*, Vol. 15, Studies in Applied Philosophy, Epistemology and Rational Ethics, Springer, Berlin-New York, doi: 10.1007/978-3-319-05107-9.
- Bonini, L., Ugolotti Serventi, F., Bruni S., Maranesi, M., Bimbi M., Simone L., Rozzi S., Ferrari P.F., Fogassi L., (2012), Selectivity for grip type and action goal in macaque inferior parietal and ventral premotor grasping neurons, *Journal of Neurophysiology,* 108 (6), pp. 1607–1619, doi: 10.1152/jn.01158.2011.
- Bonini, L., Maranesi M., Livi A., Fogassi L., and Rizzolatti G., (2014), Space-dependent representation of objects and other's action in monkey ventral premotor grasping neurons, *The Journal of Neuroscience* 34, pp. 4108–4119, doi: 10. 1523/JNEUROSCI.4187-13.2014.
- Borghi, A. M., and Riggio L., (2009), Sentence comprehension and simulation of object temporary, canonical and stable affordances, *Brain Research,* 1253, pp. 117-28, doi: 10.1016/j.brainres.2008.11.064. Epub 2008 Dec 3.
- Borghi, A. M., Gianelli C., and Scorolli C., (2010), Sentence comprehension: effectors and goals, self and others. An overview of experiments and implications for robotics, *Frontiers in Neurorobotics*, 4 (3), doi 10.3389/fnbot.2010.00003.
- Borghi, A. M., and Cimatti F., (2010), Embodied cognition and beyond: acting and sensing the body, *Neuropsychologia*, 48, pp. 763–773.
- Borghi, A.M., Flumini A., Natraj N., (2012), One hand, two objects: Emergence of affordance in contexts, *Brain and Cognition*, 80, pp. 64–73.
- Borghi, A. M., and Riggio L., (2015), Stable and variable affordances are both automatic and flexible, *Frontiers in Human Neuroscience* 9: 351, doi: 10.3389/fnhum.2015.00351. eCollection 2015.
- Borra, E., Belmalih A., Calzavara R., Gerbella M., Murata A., Rozzi S., Luppino G., (2008), Cortical Connections of the Macaque Anterior Intraparietal (AIP) Area. *Cerebral Cortex*, 18, pp. 1094–1111.
- Bosco, A., Breveglieri R., Reser D., Galletti C. and Fattori P., (2014), Multiple Representation of Reaching Space in the Medial Posterior Parietal Area V6A, *Cerebral Cortex*, doi: 10.1093/cercor/bht420.
- Bradie, M., and Brand M., (1980), *Action and Responsibility*, Bowling Green Studies in Applied Philosophy. Vol. 2, Bowling Green, Ohio.
- Brand, M., (1984), *Intending and Acting. Toward a Naturalized Action Theory.* Cambridge, MA: MIT Press.
- Bratman, M.A., (1984), Two faces of intention, *The Philosophical Review*, 93, pp. 3 – pp. 375-405.
- Bratman, M.A., (1999), *Faces of Intention*, Selected Essays on intention and agency, Cambridge University Press.
- Bratman, M.A., (2007), *Structures of Agency. Essays,* Oxford University Press.
- Bratman, M.A., (2014), *Shared Agency: A Planning Theory of Acting Together*, Oxford University Press.
- Brenet, J. P. (2014), Agent Sense in Averroes and Latin Averroism, In Silva, J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer. Switzerland, pp. 147-166.
- Briscoe, R.E., (2008), Vision, Action, and Make-Perceive, *Mind and Language*, 23, pp. 457–497.
- Briscoe, R. E., (2009), Egocentric spatial representation in action and perception,

Philosophy and Phenomenological Research, 79, pp. 423–60.

- Briscoe, R. E., (2011), Mental imagery and the varieties of amodal Perception", *Pacific Philosophical Quarterly*, 92, pp. 153-173.
- Brogaard, B., (2011), Conscious vision for action versus unconscious vision for action? *Cognitive Science*, 35, pp. 1076–104.
- Brooks, A., and Akins K., (2005), *Cognition and the Brain. The Philosophy and Neuroscience Movement*, Cambridge: Cambridge University Press.
- Bruno, N., (2001), When does action resist visual illusions? *Trends in Cognitive Sciences*, 5, pp. 385–8.
- Bruno, N., and Bernardis P., (2002), Dissociating perception and action in Kanizsa's compression illusion, *Psychonomic Bulletin & Review*, 9, pp. 723–30.
- Bruno, N. and Battaglini P. P., (2008), Integrating perception and action through cognitive neuropsychology (broadly conceived), *Cognitive Neuropshycology*, 25 (7–8), pp. 879–890.
- Bruno, N., Bernardis P., Gentilucci M., (2008), Visually guided pointing, the Müller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: Conclusions from 33 independent studies, *Neuroscience and Biobehavioral Reviews*, 32, pp. 423–437.
- Bruno, N., and Franz V. H., (2009), When is grasping affected by the Müller-Lyer illusion? A quantitative review, *Neuropsychologia,* 47, pp. 1421–1433.
- Bruzzo, A., Borghi A. M., and Ghirlanda S. (2008), Hand–object interaction in perspective, *Neuroscience Letters*, 441, pp. 61–65.
- Buccino, G., Sato M., Cattaneo L., Rodà F., and Riggio L., (2009), Broken affordances, broken objects: a TMS study, *Neuropsychologia* 47, pp. 3074–3078, doi: 10.1016/j.neuropsychologia.2009.07.003.
- Bullier, J., Hupé J. M., James A. C., and Girard P., (2001), The role of feedback connections in shaping the responses of visual cortical neurons, *Progress in Brain Research*, 134, pp. 193–204.
- Burge, T. (2010), *The Origins of Objectivity*, Oxford: Clarendon.
- Butterfill, S.A., Sinigaglia C., (2012), Intention and Motor Representation in Purposive Action", *Philosophy and Phenomenological Research*, Vol. LXXXVIII. DOI: 10.1111/j.1933-1592.2012.00604.x _ 2012.
- Caligiore, D., Borghi A. M., Parisi D., and Baldassarre G., (2010), TRoPICALS: a computational embodied neuroscience model of experiments on compatibility effects, *Psychological Review,* 117, pp. 1188–1228, doi: 10.1037/a0020887
- Caligiore, D., Borghi A. M., Parisi D., Ellis R., Cangelosi A., and Baldassarre G., (2013), How affordances associated with a distractor object affect compatibility effects: a study with the computational model TRoPICALS. *Psychological Research* 77, pp. 7–19, doi: 10.1007/s00426-012-0424-1
- Campbell, J., (2011), Tyler Burge: Origins of objectivity, *Journal of Philosophy*, 108, pp. 269–85.
- Cardellicchio, P., Sinigaglia C., and Costantini M., (2011), The space of affordances: a TMS study, *Neuropsychologia*, 49(5), pp. 1369–1372.
- Castiello, U., (2005), The Neuroscience of Grasping, *Nature Reviews*, 6 (9), pp 726-36, doi: 10.1038/nrn1744.
- Castiello, U., and Begliomini C., (2008), The cortical control of visually guided grasping, *Neuroscientist*, 14(2), pp. 157-70, doi: 10.1177/1073858407312080. Epub 2008 Jan 24.
- Chalupa, L. and Werner J., (Eds.), *The Visual Neurosciences,* Cambridge, MA: MIT

Press, pp. 1119- 1128.

- Chao, L. L., and Martin A., (2000), Representation of manipulable man-made objects in the dorsal stream, *NeuroImage*, 12, pp. 478–484.
- Charland, L., (2002), The Natural Kind Status of Emotion, *British Journal for the Philosophy of Science*, 53: pp. 511–37.
- Chemero, A. (2009), *Radical Embodied Cognitive Science*, The MIT Press, Cambridge (MA).
- Chinellato, E., and del Pobil A.P., (2015), *The Visual Neuroscience of Robotic Grasping, Achieving Sensorimotor Skills through Dorsal-Ventral Stream Integration*, Springer International Publishing Switzerland.
- Christopher, A. Buneo A. C., Andersen R. A., (2006), The posterior parietal cortex: Sensorimotor interface for the planning and online control of visually guided movements, *Neuropsychologia* 44, pp. 2594–2606.
- Cisek, P., (2007), Cortical mechanisms of action selection: the affordance competition hypothesis, *Philosophical Transactions of the Royal Society, Biological Sciences* 362, pp. 1585–1599, doi: 10.1098/rstb.2007.2054.
- Cisek, P., and Kalaska J.F., (2010), Neural Mechanisms for Interacting with a World Full of Action Choices, *Annual Review of Neuroscience*, 33, pp. 269-98.
- Clark, A., (2009), Perception, action, and experience: Unraveling the golden braid. *Neuropsychologia*, 47 (6), pp. 1460-1468. doi: 10.1016/j.neuropsychologia.2008.10.020.
- Colombetti, G., (2007), Enactive appraisal, *Phenomenology and the Cognitive Science*, 6, pp. 527–546, doi: 10.1007/s11097-007-9077-8.
- Colombetti, G., (2014), *The Feeling Body Affective Science Meets the Enactive Mind*, Cambridge, MA: The MIT Press.
- Cook, R., Bird G., Catmur C., Press C., Heyes C., (2014), Mirror neurons: From origin to function, *Behavioral and Brain Ssciences,* 37, pp. 177–241 doi: 10.1017/S0140525X13000903.
- Corcilius, K., (2014), Activity, Passivity, and Perceptual Discrimination in Aristotle, In Silva, J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 31-54.
- Costantini, M., Ambrosini E., Tieri G., Sinigaglia C., and Committeri G., (2010), Where does an object trigger an action? An investigation about affordance in space, *Experimental Brain Research*, 207, pp. 95–103.
- Costantini, M., Ambrosini E., Scorolli C., and Borghi A. M., (2011a), When objects are close to me: affordances in the peripersonal space, *Psychonomic Bullettin & Review* 18, pp. 302–308, doi: 10.3758/s13423-011-0054-4.
- Costantini, M., Committeri G., and Sinigaglia C., (2011b), Ready both to your and to my hands: mapping the action space of others, *PLoS One* 6:e17923, doi: 10. 1371/journal.pone.0017923.
- Craig, A. D., (2002), Opinion: how do you feel? Interoception: the sense of the physiological condition of the body, *Nature Reviews Neuroscience,* 3, pp. 655–666.
- Craig, A. D., (2003), Interoception: The Sense of the Physiological Condition of the Body, *Current Opinion in Neurobiology*, 13(4), pp. 500–505.
- Craighero, L., Bello A., Fadiga L., and Rizzolatti G., (2002), Hand action preparation influences the responses to hand pictures, *Neuropsychologia*, 40(5), pp. 492-502.
- Creem-Regehr, S. H., and Lee J. N., (2005), Neural representations of graspable objects: Are tools special? *Cognitive Brain Research*, 22, pp. 457–469.

- Cutting, (2003), Reconceiving perceptual space, In Hecht, H., Schwartz R., and Atherton M., (Eds.) (2003), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*, Cambridge, MA: MIT Press, pp. 215-238.
- Davidson, D., (1963/1980), Actions, reasons, and causes. In *Essays on actions and events*, Oxford: Clarendon Press.
- Davidson, D. (1980), *Essays on actions and events*, Oxford: Clarendon.
- de Sousa, R. Emotion, *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL=<http://plato.stanford.edu/archives/spr2014/entries/emotion/>.
- De Stefani E., Innocenti A., De Marco D., Busiello B., Ferri F., Costantini M., Gentilucci M., (2014), The spatial alignment effect in near and far space: a kinematic study, *Experimental Brain Research*, 232 (7), pp. 2431-2438.
- Decety, J., and Grèzes J., (2006), The power of simulation: imagining one's own and other's behavior, *Brain Research* 1079, pp. 4–14.
- Deco, G. and Lee T.S., (2004), The role of early visual cortex in visual integration: a neural model of recurrent interaction, *European Journal of Neuroscience* 20, pp. 1089–1100.
- DeLoache, J.S., Pierroutsakos S. L., Uttal D. H., Rosengren K. S. and Gottlieb A., (1998), Grasping the nature of pictures, *Psychological Science,* 9, pp. 205–210.
- Deloache, J. S., (2004), Becoming symbol-minded, *Trends in Cognitive Sciences*, 8 (2), pp. 66-70, doi: 10.1016/j.tics.2003.12.004.
- Dennis, N. A., Hermsdörfer J., (2009), *Sensorimotor Control of Grasping: Physiology and Pathophysiology*, Cambridge University Press, Cambridge.
- Derbyshire, N., Ellis R., Tucker M., (2006), The potentiation of two components of the reach-to-grasp action during object categorisation in visual memory, *Acta Psychologica*, 122 (1), pp. 74–98
- Dijkerman, H. C., McIntosh, R. D., Schindler, I., Nijboer, T. C. W., and Milner, A. D. (2009), Choosing between alternative wrist postures: Action planning needs perception. *Neuropsychologia*, doi: 10.1016/j.neuropsychologia.2008.12.002.
- Dokic, J., (2010), Perceptual Recognition and the Feeling of Presence, in Nanay B., (Ed.) (2010), *Perceiving the World*, pp. 33-54, Oxford University Press.
- Dokic, J. (2012), Seeds of Knowledge: Noetic Feelings and Metacognition, in *Foundations of Metacognition*, edited by M. Beran, J. L. Brandl, J. Perner, and J. Proust. Oxford: Oxford University Press, pp. 302–20.
- Döring, S. A., (2007), Seeing What to Do: Affective Perception and Rational Motivation, *Dialectica* 61, pp. 363-394.
- Duncan, S. and Barrett L. F., (2007), Affect as a form of cognition: a neurobiological analysis, *Cognition & Emotion*, 21, pp.1184–1211, doi: 10.1080/02699930701437931.
- Elster, J., (2003), Emotion and Action, in R. C. Solomon (Ed.), (2010), *Thinking About Feeling: Contemporary Philosophers on Emotions*, Oxford; New York: Oxford University Press.
- Fadiga, L., Fogassi L., Gallese V., Rizzolatti G., (2000), Visuomotor neurons: ambiguity of the discharge or 'motor' perception? *International Journal of Psychophysiology,* 35, pp. 165-177.
- Fattori, P., Gamberini M., Kutz D. F., and Galletti C., (2001), 'Arm-reaching' neurons in the parietal area V6A of the macaque monkey, *European Journal of Neuroscience,* 13, pp. 2309–2313, doi: 10.1046/j.0953-816x.2001.01618.x.
- Fattori, P., Kutz D. F., Breveglieri R., Marzocchi N., and Galletti C., (2005), Spatial

tuning of reaching activity in the medial parieto-occipital cortex (area V6A) of macaque monkey, *European Journal of Neuroscience,* 22, pp. 956–972, doi: 10.1111/j.1460- 9568.2005.04288.x.

- Feagin, S. L., (1998), Presentation and representation, *Journal of Aesthetics and Art Criticism*, 56, pp. 234–240.
- Fearraina, S., Battaglia-Mayer A., Genovesio A., Archambault P., and Caminiti R., (2009), Parietal encoding of action in depth, *Neuropsychologia*, doi: 10.1016/j.neuropsychologia.2008.12.028.
- Ferretti, G., (forthcoming), Neurophysiological States and Perceptual Representations: The case of Action Properties detected by the Ventro-Dorsal Visual Stream, In Magnani L., and Casadio C. (Eds.), *Model-Based Reasoning in Science and Technology. Models and Inferences: Logical, Epistemological, and Cognitive Issues*, Series "Sapere", *Studies in Applied Philosophy and Rational Ethics,* Springer.
- Ferretti, G., and Alai M., (forthcoming), Enactivism, Representations and Canonical Neurons, *Argumenta, 1 (2).*
- Ffytche, D. H., Guy C. N., and Zeki S., (1995), The parallel visual motion inputs into areas V1 and V5 of human cerebral cortex. *Brain* 118, pp. 1375–1394. doi: 10.1093/brain/118. 6.1375.
- Findlay J., M., and Gilchrist I. D., (2003), *Active vision: the psychology of looking and seeing*, Oxford University Press, Oxford.
- Fluet, M. C., Baumann M. A., Scherberger H., (2010), Context-specific grasp movement representation in macaque ventral premotor cortex, *The Journal of Neuroscience*, 30: pp. 15175-15184.
- Fogassi, L., Gallese V., Fadiga L., Luppino G., Matelli M., and Rizzolatti G., (1996), Coding of peripersonal space in inferior premotor cortex (area F4), *Journal of Neurophysiology,* 76, pp. 141–157.
- Fogassi, L., and Luppino G., (2005), Motor functions of the parietal lobe, *Current Opinion in Neurobiology*, 15: pp. 626–631, doi: 10.1016/j.conb.2005.10.015.
- Foxe, J. J., and Simpson G. V., (2002), Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing, *Experimental Brain Research*, 142, pp. 139–150, doi: 10.1007/s00221-001-0906-7.
- Franz, V. H., and Gegenfurtner K. R., (2008), Grasping visual illusions: Consistent data and no dissociation, *Cognitive Neuropsychology*, 25 (7–8), pp. 920–950.
- Franz, V.H., Hesse C., Kollath S., (2009), Visual illusions, delayed grasping, and memory: No shift from dorsal to ventral control, *Neuropsychologia*, 47, pp. 1518–1531.
- Fridland, E., (2012), Nivedita Gangopadhyay, Michael Madary, and Finn Spencer (Eds.), Perception, action, and consciousness: sensorimotor dynamics and the two visual systems. Book Review, *Phenomenology and the Cognitive Science*, doi: 10.1007/s11097-012-9260-4.
- Frijda, N., (1986), *The Emotions*, Cambridge: Cambridge University Press.
- Gallese, V., (2000), The inner sense of action. Agency and motor representations, *Journal of Consciousness Studies,* 7 (10), pp. 23-40.
- Gallese, V., (2005), Embodied simulation: from neurons to phenomenal experience, Phenomenology and the Cognitive Science, 4, pp. 23–48.
- Gallese, V., (2007), The "Conscious" Dorsal Stream: Embodied Simulation and its Role in Space and Action Conscious Awareness, *Psyche,* 13 (1), pp. 1-20.
- Gallese, V., (2009), Motor abstraction: a neuroscientic account of how action goals and intentions are mapped and understood, *Psychological Research*, 73, pp. 486–498.

- Gallese, V., Craighero L., Fadiga L. and Fogassi L., (1999), Perception Through Action, *Psyche*, 5 (21).
- Gallese, V., and Metzinger T., (2003), Motor Ontology. The representational reality of goals, actions and selves, *Philosophical Psychology*, 16(3), pp. 365-388.
- Gallese, V., and Sinigaglia C., (2011a), How the body in action shapes the self, *Journal of Consciousness Studies*, 18 (7-8), pp. 117-143.
- Gallese, V. and Sinigaglia C., (2011b), What is so special about embodied simulation? *Trends in Cognitive Sciences*, 15 (11), pp. 512-519.
- Gallese, V., and Sinigaglia, C. (2011c), Cognition in action: A new look at the motor system. In Metcalfe J., H. Terrace (Ed.), (2011), *Joint Attention and Agency.* Oxford: Oxford University Press.
- Gallese, V. and Sinigaglia C., (2013) Cognition in action: A new look at the cortical motor system, In Metcalfe J. and Terrace H. S., (Eds.) Agency and Joint Attention, Oxford University Press, New York, pp. 178-195.
- Gallese, V. and Sinigaglia C., (2014), Understanding action within the motor system, *Behavioral and Brain Science,* 37 (2), pp. 199-200, doi: 10.1017/S0140525X13002288.
- Gangopadhyay, N., and Kiverstein J., (2009), Enactivism and the Unity of Perception and Action, *Topoi,* 28 (1), pp. 63-73.
- Gangopadhyay, N., (2010), Experiential blindness revisited. In defense of a case of embodied cognition, *Cognitive Systems Research*, 11, pp. 396–407.
- Gentilucci, M., Daprati E., Toni I., Chieffi S., Saetti M.C., (1995), Unconscious updating of grasp motor program. Experimental Brain Research, 105 (2), pp. 291-303.
- Gerlach, C., Law I., Paulson O.B., (2002), When action turns into words. Activation of motor-based knowledge during categorization of manipulable objects, *Journal of Cognitive Neuroscience*, 14, pp. 1230–1239.
- Gibson, J. J., (1966), *The senses considered as perceptual systems,* Houghton Mifflin.
- Gibson, J. J., (1979), *The ecological approach to visual perception,* Boston: Houghton Mifflin.
- Girardi, G., Lindemann O., and Bekkering H., (2010), Context effects on the processing of action-relevant object features *Journal of Experimental Psychology. Human Perception and Performance*, 36, pp. 330–340, doi: 10.1037/a0017180
- Goldie, P., (2000), *The Emotions: A Philosophical Exploration*, Oxford: Oxford University Press.
- Gonzalez, C., Ganel T., Whitwell R., Morrissey B., and Goodale M., (2008), Practice makes perfect, but only with the right hand: Sensitivity to perceptual illusions with awkward grasps decreases with practice in the right, but not the left hand, *Neuropsychologia*, 46, pp. 624-631.
- Goodale, M.A., Milner A.D., (1992), Separate visual pathways for perception and action, *Trends in Neurosciences*, 15, pp. 20–5.
- Grade, G., Pesenti M., Edwards M.G., (2015), Evidence for the embodiment of space perception: concurrent hand but not harm action moderates reachability and egocentric distance perception, *Frontiers in Psychology*, doi: 10.3389/fpsyg.2015.00862.
- Grafton, S. T., Fadiga L., Arbib M. A., and Rizzolatti G., (1997), Premotor cortex activation during observation and naming of familiar tools, *NeuroImage*, 6, pp. 231–236.
- Grafton, S. T., and Hamilton A. F. D. C., (2007). Evidence for a distributed

hierarchy of action representation in the brain, *Human Movement Science,* 26, pp. 590–616, doi: 10.1016/j.humov.2007.05.009.

- Graziano, M., (2009), *The Intelligent Movement Machine: An Ethological Perspective on the Primate Motor System*, Oxford: Oxford University Press.
- Greenspan, P., (1988), *Emotions and Reasons: an Inquiry into Emotional Justification*, New York: Routledge, Chapman and Hall.
- Greenspan, P., (2000), Emotional Strategies and Rationality, *Ethics*, 110, pp. 469–487.
- Grezes, J., and Decety J., (2002), Does visual perception of object afford action? Evidence from a neuroimaging study, *Neuropsychologia*, 40, pp. 212–222.
- Grill-Spector, K., Kourtzi Z., and Kanwisher N., (2001), The lateral occipital complex and its role in object recognition, *Vision Research*, 41, pp. 1409-1422.
- Grush, R., (1998), Skill and Spatial Content, *Electronic Journal of Analytic Philosophy,* 6.
- Grush, R., (2000), Self, World, and Space: The Meaning and Mechanisms of Ego- and Allocentric Spatial Representation, *Brain and Mind,* 1, pp. 59–92.
- Grush, R., (2007), Skill Theory v2.0: Dispositions, Emulation, and Spatial Perception, *Synthese* 159, pp. 389–416.
- Gutteling, T. P., Park S., Kenemans J., Neggers S. F., (2013), TMS of the anterior intraparietal area selectively modulates orientation change detection during action preparation, *Journal of Neurophysiology*, 110, pp. 33–41.
- Gutteling, T. P., Petridou N., Dumoulin S., O., Harvey B., M., Aarnoutse E. J., Kenemans J. L., Neggers S.F.W. (2015), Action Preparation Shapes Processing in Early Visual Cortex, *The Journal of Neuroscience*, 35 (16), pp. 6472-6480; doi: 10.1523/JNEUROSCI.1358-14.2015.
- Hadjidimitrakis, K., Breveglieri R., Placenti G., Bosco A., Sabatini S. P. and Fattori P., (2011), Fix your eyes in the space you could reach: neurons in the macaque medial parietal cortex prefer gaze positions in peripersonal space. *PLoS ONE,* doi: 10.1371/journal.pone.0023335.
- Hadjidimitrakis, K., Bertozzi F., Breveglieri R., Fattori P., and Galletti C., (2013), Body-centered, mixed, but not hand-centered coding of visual targets in the medial posterior parietal cortex during reaches in 3D space, *Cerebral Cortex*, 4, (12), pp. 3209-3220, doi: 10.1093/cercor/bht181.
- Hatfield, G., (2014), Activity and Passivity in Theories of Perception: Descartes to Kant. In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 275-290.
- Hecht, H., Schwartz R., and Atherton M., (Eds.), (2003), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space.* Cambridge, MA: MIT Press, pp. 215-238.
- Helbig, H. B., Graf M., Kiefer M., (2006), The role of action representations in visual object recognition, *Experimental Brain Research*, 174: pp. 221–228, doi: 10.1007/s00221-006-0443-5.
- Holmes, N.P., and Spence C., (2004), The body schema and multisensory representation(s) of peripersonal space, *Cognitive Processing*, 5, pp. 94–105, doi: 10.1007/s10339-004-0013-3.
- Hommel, B., Müsseler J., Aschersleben G., and Prinz W., (2001), The theory of event coding: A framework for perception and action planning, *Behavioral and Brain Sciences*, 24, pp. 849–931.
- Howard, I. P., and Rogers B. J., (2002), *Seeing in depth: Vol. 2. Depth perception.*

Thornhill, Ontario, Canada: University of Toronto Press.

- Hurley, S. L., (1998), *Consciousness in Action*, Cambridge, MA: Harvard University Press.
- Hursthouse, R., (1991), Arational Actions, *The Journal of Philosophy*, 88 (2), pp. 57-68.
- Husserl, E., (1907) *Thing and Space: Lectures of 1907,* Husserliana: Edmund Husserl - Collected Works, Rojcewicz R. (Trans.) (1998).
- Hutto, D. D., (2005), Knowing What? Radical versus Conservative Enactivism, *Phenomenology and the Cognitive Sciences*, 4 (4), pp. 389-405.
- Hutto, D. D., Myin E., (2013), *Radicalizing Enactivism: Basic Minds without Content.* The MIT Press, Cambridge (MA).
- Iachini, T., Borghi A.M., Senese V.P., (2008), Categorization and sensorimotor interaction with objects, *Brain and Cognition*, 67, pp. 31-43.
- Jackson, S. J., Newport R. W., Husain M., Fowlie J., O'Donoghue M. and Bajaj N., (2009), There may be more to reaching than meets the eye: Re-thinking optic ataxia, *Neuropsychologia*, 47, pp. 1397-1408, doi: 10.1016/j.neuropsychologia.2009.01.035.
- Jacob, P., Jeannerod M., (2003), *Ways of seeing. The Scope and Limits of Visual Cognition*, Oxford University Press: Oxford.
- Jacob, P., (2005), Grasping and Perceiving Objects, In Brooks A. and Akins K., (Eds.) (2005), *Cognition and the Brain. The Philosophy and Neuroscience Movement*, Cambridge: Cambridge University Press, pp. 252- 283.
- Jacob, P., (2008), The scope and limits of enactive approaches to visual experience, *Perception*, 37 (3), pp. 446-461.
- Jacob, P., (2014), The legacy of Marc Jeannerod, *Neuropsychologia*, 53, 3-5.
- Jacob, P., (2015), Action based accounts (of perception), In Matthen M. (Ed.) (2015), *Handbook of the Philosophy of Perception*, Oxford, Oxford University Press, pp. 217-236.
- Jacob, P., and de Vignemont F., (2010), Spatial coordinates and phenomenology in the two-visual systems model. In Gangopadhyay N., Madary M. and Spicer F. (Eds.), *Perception, Action and Consciousness*, Oxford, Oxford University Press, pp. 125-144.
- James, T., Humphrey G., Gati J., Menon R., and Goodale M., (2002), Differential effects of viewpoint on object-driven activation in dorsal and ventral stream, *Neuron*, 35, pp. 793-801.
- Janssen, P., and Scherberger H., (2015), Visual guidance in control of grasping, *Annual Review of Neuroscience*, 8 (38), pp. 69-86, doi: 10.1146/annurev-neuro-071714-034028.
- Janyan, A., and Slavcheva G. V., (2012), When left feels right: asymmetry in the affordance effect, *Cognitive Processing*, Suppl. 1, pp. 199-202, doi: 10.1007/s10339-012-0450-3.
- Jax, S. A., and Buxbaum L. J., (2010), Response interference between functional and structural actions linked to the same familiar object, *Cognition,* 115(2), pp. 350–355, doi: 10.1016/j.cognition.2010.01.004.
- Jeannerod, M., (1994), The representing brain: Neural correlates of motor intention and imagery, *Behavioral and Brain Sciences*, 17, pp. 187–245.
- Jeannerod, M., (1997), *The Cognitive Neuroscience of Action*, Oxford: Blackwell.
- Jeannerod, M., (2006), *Motor Cognition: What Actions Tell the Self*, Oxford: Oxford University Press.
- Jeannerod, M. and Jacob P., (2005), Visual cognition: A new look at the two-visual

systems model, *Neuropsychologia*, (43), pp. 301–12.

- Kalénine, S., Shapiro A.D., Flumini A., Borghi A.M., Buxbaum L.J., (2013), Visual context modulates potentiation of grasp types during semantic object categorization, *Psychonomic Bullettin & Review*, doi: 10.3758/s13423-013-0536-7.
- Kandel, E. R., Schwartz J. H., Jessell T. M., Siegelbaum S. A., Hudspeth A. J., (2013), *Principles of Neural Science*, McGrawHill Medical.
- Kaukua, J., (2014), Avicenna on the Soul's Activity in Perception, In Silva, J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 99-116.
- Kellenbach, M.L., M. Brett, K. Patterson, (2003), The importance of manipulability and action in tool representation, *Journal of Cognitive Neuroscience*, 15, pp. 30–46.
- Kelly, S. D., (2008), Content and Constancy: Phenomenology, Psychology, and the Content of Perception, *Philosophy and Phenomenological Research*, 76 (3), pp. 682–690, doi: 10.1111/j.1933-1592.2008.00164.x.
- Kenny, A., (2003), *Action, Emotion and Will*, Routledge. Taylor and Francis Group.
- Khalsa, S. S., Rudrauf D., Feinstein J.S., and Tranel. D., (2009), The Pathways of Interoceptive Awareness, *Nature Neuroscience* 12 (12), pp. 1494–1496.
- Kitadono, K. and Humphreys G. W., (2009), Sustained interactions between perception and action in visual extinction and neglect: Evidence from sequential pointing, *Neuropsychologia*, 47 (6), pp. 1592, 1599, doi: 10.1016/j.neuropsychologia.2008.11.010.
- Klistorner, A., Crewther D. P., and Crewther S. G., (1997), Separate Magnocellular and Parvocellular contributions from Temporal Analysis of the Multifocal VEP, *Vision Research*, 37, pp. 2161–2169, doi: 10.1016/S0042-6989(97)00003-5.
- Knight, R. T., Staines W. R., Swickc D., and Chaoc L. L., (1999), Prefrontal cortex regulates inhibition and excitation in distributed neural networks, *Acta Psychologica*, 101, pp. 159–178.
- Koenderink, J. J., and van Doorn A. J., (2003), Pictorial space, in Hecht, H., Schwartz, R., & Atherton, M. (eds.) (2003), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*, Cambridge, MA: MIT Press, pp. 215-238.
- Koenderink, J. J., van Doorn A. J., and Wagemans J (2011), Depth. *i-Perception*, 2, pp. 541-564.
- Kravitz, D. J., Saleem K. I., Baker C. I., and Mishkin M., (2011), A new neural framework for visuospatial processing, *Nature Reviews Neuroscience*, 12, pp. 217–30.
- Kravitz, D. J., K. S. Saleem, C. I. Baker, L. G. Ungerleider, and Mishkin. M., (2013), The ventral visual pathway: an expanded neural framework for the processing of object quality, *Trends in Cognitive Sciences*, 17 (1), pp. 26-49.
- Kulvicki, J., (2014), *Images*, New York: Routledge.
- Lagerlund, H., (Ed.), (2007), *Forming the Mind. Essays on the Internal Senses and the Mind/Body Problem from Avicenna to the Medical Enlightenment*, Springer Studies in the History of Philosophy of Mind, Vol. 5. The Netherlands.
- Lähteenmäki, V., (2014) Locke and Active Perception, In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy* Springer, Switzerland, pp. 223-240.
- Laylock, R., Crewther S. G., and Crewther D. P., (2007), A role for the "magnocellular advantage" in visual impairments in neurodevelopmental and psychiatric disorders, *Neurosciesnce and Biobehavioral Reviews,* 31, pp. 363–376. doi:10.1016/ j.neubiorev.2006.10.003.
- Lazarus, R.S., (1991), *Emotion and Adaptation*, New York & Oxford: Oxford

University Press.

- Ledoux, J., (1996), *The emotional brain: the mysterious underpinnings of emotional life*, New York: Simon and Schuster.
- Leijenhorst, C., (2014), Active Perception from Nicholas of Cusa to Thomas Hobbes, In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy* Springer, Switzerland, pp. 167-186.
- Levinson, J., (1998), Wollheim on pictorial representation, *Journal of Aesthetics and Art Criticism*, 56, pp. 227–233.
- Lopes, D. M., (1996), *Understanding pictures*, Oxford: Oxford University Press.
- Lopes, D. M., (2005), *Sight and Sensibility: Evaluating Pictures*, Oxford University Press.
- Luppino, G., Murata A., Govoni P. and Matelli M., (1999), Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4), *Experimental Brain Research*, (128), pp. 181-187.
- MacKay, D. M., (1962), Theoretical models of space perception. In Muses C. A. (1962) (Ed.), *Aspects of the theory of artificial intelligence,* Plenum Press.
- MacKay, D. M., (1967), Ways of looking at perception. In W. Wathen-Dunn (Ed.) (1967) *Models for the perception of speech and visual form,* The MIT Press.
- MacKay, D. M., (1973), Visual stability and voluntary eye movements. In Jung. R. (Ed.), *Handbook of sensory physiology, vol. VII/3A,* Springer.
- Maranesi, M., Bonini L., and Fogassi L., (2014), Cortical processing of object affordances for self and others' action, *Frontiers in Psychology,* 5, 538, doi: 10.3389/fpsyg. 2014.00538.
- Marotta, J., and Goodale M., (2001), The role of familiar size in the control of grasping, *Journal of Cognitive Neuroscience*, 13, pp. 8-17.
- Martin, J. R., and Dokic J., (2013), Seeing Absence or Absence of Seeing? *Thought: A Journal of Philosophy,* 2 (2), pp. 117–125.
- Matelli, M., Luppino G., Govoni P., and Geyer S., (1996), Anatomical and functional subdivisions of inferior area 6 in macaque monkey, *Society for Neuroscience, Abstracts,* 22 (3), pp. 2024–2024.
- Matthen, M., (2005), *Seeing, doing and knowing: a philosophical theory of sense perception*, Oxford: Oxford University Press.
- Matthen, M., (2010), Two Visual Systems and the feeling of presence, pp. 107-125, in Nivedita Gangopadhyay, Michael Madary, and Finn Spencer (Eds.), (2010), *Perception, action, and consciousness: sensorimotor dynamics and the two visual systems*, Oxford University Press.
- Mausfeld, R., (2003), Conjoint representations, In H. Hecht, R. Schwartz, and M. Atherton (Eds.), *Looking into pictures: An interdisciplinary approach to pictorial space*, Cambridge, MA: MIT Press, pp. 17–60.
- McIntosh, R. D., and Lashleya G., (2008), Matching boxes: Familiar size influences action programming, *Neuropsychologia*, 46, pp. 2441–4.
- McIntosh, R.D, Schenk, T., (2009), Two visual streams for perception and action: current trends, *Neuropsychologia,* 47 (6), pp. 1391-1396, doi: 10.1016/j.neuropsychologia.2009.02.009, Epub 2009 Feb 13.
- Mele, A. R., (1987), *Irrationality. An Essay on Akrasia, Self-Deception and Self-Control*, Oxford University Press.
- Mele, A. R., (1992), *Springs of Action. Understanding Intentional Behavior*, Oxford

University Press.

- Mele, A.R., (1995), *Autonomous Agents. From Self-Control to Autonomy*, Oxford Unviersity Press.
- Mele, A. R., (2003), *Motivation and Agency*, Oxford University Press.
- Mele, A.R., (2009), *Effective Intentions. The Power of Conscious Will,* Oxford Unviersity Press.
- Mele, A.R., (2014), *Free. Why Science Hasn't Disproved Free Will, Oxford University Press.*
- Merleau-Ponty, M., (1962), *The Phenomenology of Perception*, Smith, C. (trans.), London: Routledge and Kegan Paul.
- Milner, A. D., and Goodale M. A., (1995/2006), *The Visual Brain in Action*, 2nd edition, Oxford: Oxford University Press.
- Milner, A. D., and Goodale M. A., (2010), Cortical Visual Systems for Perception and Action, In Gangopadhyay, Nivedita, Michael Madary, Finn Spencer (eds.) (2010), *Perception, Action, and Consciousness: Sensorimotor Dynamics and the Two Visual Systems*, Oxford: Oxford University Press, pp. 71-95.
- Milner, A. D., and Goodale M. A., (2008), Two visual systems re-viewed, *Neuropsychologia*, 46, pp. 774–85.
- Morgado, N., Muller D., Gentaz E., and Palluel-Germain R., (2011), Close to me? The influence of affective closeness on space perception, *Perception*, 40(7), pp. 877–9.
- Munakata, Y., Herd S. A., Chatham C. H., Depue B. E., Banich M. T., and O'Reilly R. C, (2011), A unified framework for inhibitory control, *Trends in Cognitive Sciences*, 15(10), pp. 453–459.
- Murata, A., Fadiga L., Fogassi L., Gallese V., Raos V., and Rizzolatti G., (1997), Object representation in the ventral premotor cortex (area F5) of the monkey, *Journal of Neurophysiology,* 78: pp. 2226-2230.
- Murata, A., Gallese V., Luppino G., Luppino G., Kaseda M., Sakata H., (2000), Selectivity for the shape, size and orientation of objects for grasping in neurons of monkey parietal area AIP, *Journal of Neurophisiology,* 79, pp. 2580-2601.
- Nanay, B., (2005), Is twofoldness necessary for representational seeing? *British Journal of Aesthetics*, 45, pp. 263–272.
- Nanay, B, (2010 a), Perception and Imagination: Amodal Perception as Mental Imagery, *Philosophical Studies,* 150, pp. 239–254.
- Nanay, B., (2010 b), Inflected and uninflected experience of pictures. In C. Abell and K. Bantinaki (Eds.), *Philosophical perspectives on depiction,* pp. 181–207, Oxford: Oxford U.P.
- Nanay, B., (Ed.), (2010 c), *Perceiving the World*, New York: Oxford University Press.
- Nanay, B., (2011 a), Perceiving pictures, *Phenomenology and the Cognitive Sciences*, 10, pp. 461–480.
- Nanay, B., (2011 b), Do we sense modalities with our sense modalities? *Ratio*, 24, pp. 299–310.
- Nanay, B., (2011 c), Action Oriented-Perception, *European Journal of Philosophy*, 20 (3), pp. 430–446, doi: 10.1111/j.1468-0378.2010.00407.x.
- Nanay, B., (2012), Perceiving tropes, *Erkenntnis,* 77: pp. 1-14.
- Nanay, B., (2013 a), Is action-guiding vision cognitively impenetrable? In: *Proceedings of the 35th Annual Conference of the Cognitive Science Society (CogSci 2013).* Hillsdale, NJ: Lawrence Erlbaum, pp. 1055-1060.
- Nanay, B., (2013 b), *Between Perception and Action*, Oxford: Oxford University Press.

- Nanay, B., (2014 a), Every act an animal act: Naturalizing action theory, In Sprevak M. and Kallestrup J., (Eds.) (2014): *New Waves in the Philosophy of Mind.* Palgrave Macmillan, pp. 226-241.
- Nanay, B., (2014 b), Empirical problems with anti-representationalism. In Brogaard B., (Ed.), (2014), *Does Perception have Content?* New York: Oxford University Press.
- Nanay, B., (2015), Trompe l'oeil and the Dorsal/Ventral Account of Picture Perception, *Review of Philosophy and Psychology*, 6, pp. 181–197, doi: 10.1007/s13164-014-0219-y.
- Nelissen, K., Joly O., Durand J. B., Todd J. T., Vanduffel W., and Orban G. A., (2009), The extraction of depth structure from shading and texture in the macaque brain, *PloS one*, *4*(12), e8306.
- Niederée, R., and Heyer D., (2003), The Dual Nature of Picture Perception: A Challenge to Current General Accounts of Visual Perception, In Hecht, H., Schwartz R., and Atherton M., (Eds.), (2003), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*, Cambridge, MA: MIT Press, pp. 215-238.
- Nisbett, R., and Ross L., (1980), *Human Inference: Strategies and Shortcomings of Social Judgment,* Englewood Cliffs, N. J. Prentice-Hall.
- Noë, A., (2008b), Précis of *Action In Perception. Philosophy and Phenomenological Research*, 76 (3), pp. 660–665, doi: 10.1111/j.1933-1592.2008.00161.x.
- Noë, A., (2004), *Action in Perception.* The MIT Press, Cambridge (MA).
- Noë, A., (2008a), Reply to Campbell, Martin, and Kelly, *Philosophy and Phenomenological Research*, 76 (3), pp. 691–706, doi: 10.1111/j.1933-1592.2008.00165.x.
- Noë, A., (2012), *Varieties of Presence*, Harvard University Press.
- Nowak, D. A., and Hermsdörfer J., (2009), *Sensorimotor Control of Grasping: Physiology and Pathophysiology*, Cambridge University Press.
- O'Regan, K J., Noë A., (2001), A sensorimotor account of vision and visual consciousness, *Behavioral and the Brain Science,* n. 24, pp. 983-984.
- Ohman, A., Flykt A., and Esteves F., (2001), Emotion drives attention: Detecting the snake in thegrass, *Journal of Experimental Psychology General,* 130, pp. 466–478, doi: 10.1037//0096- 3445.130.3.466.
- Pacherie, E., (2000), The Content of Intentions, *Mind & Language*, 15(4), pp. 400–432.
- Pacherie, E., (2002), The Role of Emotions in the Explanation of Action, *European Review of Philosophy*, CSLI Publications, 5, p. 55-90.
- Pacherie, E., (2006), Towards a Dynamic Theory of Intentions, In Pockett, S. W.P. Banks and Gallagher S., (Eds.), (2006) *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition.* Cambridge, MA: MIT Press, pp. 145-167.
- Pacherie, E., (2007), The Sense of Control and the Sense of Agency, *Psyche*, 13 (1), pp. 1-30.
- Pacherie, E., (2008), The phenomenology of action: A conceptual framework, *Cognition*, 107 (1), pp. 179–217.
- Pacherie, E., (2011), Non-Conceptual Representations for Action and the Limits of Intentional Control, *Social Psychology*, 42, pp. 67-73, doi: 10.1027/1864-9335/a000044.
- Pani, P., Theys T., Romero M. C., and Janssen P., (2014), Grasping Execution and Grasping Observation Activity of Single Neurons in the Macaque Anterior Intraparietal Area, *Journal of Cognitive Neuroscience* 26(10), pp. 2342–2355, doi:

10.1162/jocn_a_00647.

- Paukkonen, V., (2014), Berkeley and Activity in Visual Perception, In Silva, J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 255-274.
- Paulus, W., Korinth S., Wischer S., and Tergau F., (1999), Differential inhibition of chromatic and achromatic perception by transcranial magnetic stimulation of the human visual cortex, *Neuroreport,* 10, pp. 1245–1248, doi: 10. 1097/00001756-199904260-00017.
- Pautz, A., (2010), An argument for the intentional view of visual experience, In Nanay B., (Ed.) (2010), *Perceiving the World*, New York: Oxford University Press.
- Peacocke, C., (2001), Does Perception have a Nonconceptual Content? *Journal of Philosophy*, 98, pp. 239–264.
- Pearson, J., and Westbrook F., (2015), Phantom Perception: Voluntary and Involuntary Non retinal vision, *Cell Press*, http://dx.doi.org/10.1016/j.tics.2015.03.004.
- Pesaran, B., Nelson M., and R. Andersen, (2006), 'Dorsal Premotor Neurons Encode the Relative Position of the Hand, Eye, and Goal During Reach Planning, *Neuron*, 51: pp. 125–134.
- Pessoa, L., (2008), On the relationship between emotion and cognition, *Nature Reviews Neuroscience*, 2, pp. 148–158, doi: 10.1038/ nrn2317.
- Pezzulo, G., Barca L., Bocconi A. L., Borghi A. M., (2010), When affordances climb into your mind: Advantages of motor simulation in a memory task performed by novice and expert rock climbers, *Brain and Cognition,* 73, pp. 68–73.
- Pierroutsakos, S. L., and DeLoache J. S., (2003), Infants' manual exploration of pictured objects varying in realism, *Infancy,* 4, pp. 141–156.
- Ponce, C. R., and Born R. T., (2008), Stereopsis, *Current Biology,* 18(18), R845–R850, doi: 10.1016/j.cub.2008.07.006.
- Price, J. L., (2007), Connections of orbital cortex, In Zald D. H., and Rauch S. L., (Eds.), (2007), *The Orbitofrontal Cortex*, New York, NY: Oxford University Press, pp. 38–56.
- Prinz, W., (1984), Modes of linkage between perception and action, In Prinz W., and Sanders A. F., (Eds.) (1984), *Cognition and Motor Processes,* pp. 185-193, Berlin: Springer.
- Prinz, W., (1987), Ideo-motor action, In Heuer H. and Sanders A. F. (Eds.), (1987), *Perspectives on Perception and Action*, Hillsdale, NJ: Erlbaum, pp. 47–76.
- Prinz, W., (1990), A common coding approach to perception and action, In Neumann O., and Prinz W. (Eds.), (1990), *Relationships between perception and action: Current approaches*, Berlin: Springer, pp. 167–201.
- Prinz, W., (1997), Perception and Action Planning, *European Journal of Cognitive Psychology*, 9, pp. 129-154.
- Prinz W., Beisert M., and Herwig A., (2013), *Action Science. Foundations of an Emerging Discipline*, Cambridge, MA: MIT Press.
- Prinz, J., (2004), *Gut Reactions: a Perceptual Theory of Emotion*, Oxford: Oxford University Press.
- Prosser, S., (2011), Affordances and the phenomenal character in spatial perception, *Philosophical Review*, 120, pp., 475–513.
- Proverbio, M. A., Adorni R. and D'Aniello G. E., (2011), 250 ms to code for action affordance during observation of manipulable objects, *Neuropsychologia*, 49, pp. 2711–2719.

- Pulvermüller, F. and Hauk O., (2005), Functional links between motor and language systems, *European Journal of Neuroscience*, 21, pp. 793–7.
- Pulvermüller, F., (2013), How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics, *Trends in Cognitive Sciences,* 17 (9), pp. 458-470, doi: http://dx.doi.org/10.1016/j.tics.2013.06.004.
- Ranzini, M., Borghi A.M., Nicoletti, R., (2011), With hands I do not centre! Action- and object-related effects of hand-cueing in the line bisection, *Neuropsychologia*, 49, pp. 2918–2928.
- Raos, V., Umilta M.A., Murata A, Fogassi L., Gallese V., (2006), Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey, *Journal of Neurophysiology,* 95, pp. 709-729.
- Remes, P., (2014), Plato: Interaction Between the External Body and the Perceiver in the *Timaeus*, In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosoph,* Springer, Switzerland, pp. 9-30.
- Rice, N. J., Valyear K. F., Goodale M. A., Milner A. D., and Culham J. C., (2007), Orientation sensitivity to graspable objects: an fMRI adaptation study, *Neuroimage*, 36, pp. 87-T93.
- Rietveld, E., and Kiverstein J., (2014), A Rich Landscape of Affordances, *Ecological Psychology*, 26 (4), pp. 325-352.
- Riggio, L. Patteri I., Oppo A., Buccino G., and Umiltà C., (2006), The role of affordances in inhibition of return, *Psychonomic Bullettin & Review* 13, pp. 1085–1090, doi: 10. 3758/bf03213930.
- Rizzolatti, G., (2001), The cortical motor system, *Neuron,* 31, pp. 889–901.
- Rizzolatti, G., and Camarda R., Fogassi L., Gentilucci M., Luppino G., and Matelli M., (1988), Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements, *Experimental Brain Research,* 71, pp. 491–507.
- Rizzolatti, G., and Matelli M., (2003), Two different streams form the dorsal visual system: Anatomy and functions, *Experimental Brain Research*, 153, pp. 146–57.
- Rizzolatti, G., and Sinigaglia C., (2008), *Mirrors in the brain, How our minds share actions and emotions*, Oxford University Press.
- Rizzolatti, G., and Sinigaglia C., (2013), Understanding action from the inside. In Prinz, W., Beisert M. and Herwig A., (Eds.) *Action Science: Foundations of an Emerging Discipline.* MIT Press, Cambdridge Mass, pp. 201-227.
- Roberts, R. C., (2003), *Emotions: An Essay in Aid of Moral Psychology*, New York: Cambridge University Press.
- Rogers, S., (2003), Truth and meaning in pictorial space, in Hecht, H., Schwartz R., and Atherton M., (eds.) (2003), *Looking into Pictures: An Interdisciplinary Approach to Pictorial Space*, Cambridge, MA: MIT Press, pp. 215-238.
- Romero, M.C., Van Dromme I., Janssen P., (2012), Responses to two- dimensional shapes in the macaque anterior intraparietal area, *European Journal of Neuroscience*, 36, pp. 2324 –2334.
- Romero, M. C., Van Dromme I., and Janssen P., (2013), The role of binocular disparity in stereoscopic images of objects in the macaque anterior intraparietal area, *PLoS ONE,* doi: 10.1371/journal.pone. 0055340.
- Romero, M. C., Pani P., Janssen P., (2014), Coding of Shape Features in the Macaque Anterior Intraparietal Area, *The Journal of Neuroscience*, 34 (11), pp. 4006–4021.

- Sacks, O., (2012), *Hallucinations*, Vintage.
- Sakata, H., Taira M., Murata A., and Mine S., (1995), Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey, *Cerebral Cortex*, 5 (5), pp. 429-438.
- Sakata, H., Tsutsui K., and Taira M., (2003), Representation of the 3D world in art and in the brain, *International Congress Series*, 1250, pp. 5-35.
- Sakreida, K., Menz M.M., Thill S., Rottschy C., Eickhoff B., Borghi A.M., Ziemke T., and Binkofski F., (2013), Neural pathways of stable and variable affordances: a coordinate-based meta-analysis, *F1000Posters*, Poster Number 3762.
- Schenk, T. and McIntosh R. D., (2010), Do we have independent visual streams for perception and action? *Cognitive Neuroscience*, 1, pp. 52–78.
- Schwenkler, J. L., (2014), Vision, Self-Location, and the Phenomenology of the 'Point of View, *Noûs,* 48 (1), pp. 137-155.
- Seth, A. K., (2014), A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia, *Cognitive Neuroscience*, 5 (2), pp. 97-118, doi: 10.1080/17588928.2013.877880.
- Seth, A.K., Suzuki K., and Critchley H. D., (2012), An interoceptive predictive coding model of conscious presence, *Frontiers in Psychology*, 2, doi: 10.3389/fpsyg.2011.00395.
- Shepherd, J., (2014), The Contours of Control, *Philosophical Studies,* 170 (3), pp 395-411.
- Shepherd, J., (2015a), Conscious Action/Zombie Action, *Noûs*, doi: 10.1111/nous.12086.
- Shepherd, J., (2015b), Conscious Control over Action, *Mind & Language,* 30(3), pp. 320-344.
- Shikata, E., Hamzei F., Glauche V., Koch M., Weiller C., Binkofski F., Büchel C., (2003), Functional properties and interaction of the anterior and posterior intraparietal areas in humans, *European Journal of Neuroscience*, 17 (5), pp. 1105–1110.
- Siegel, S., (2006), Which properties are represented in perception? In Gendler T. S. and Hawthorne J., (Eds.) (2006), *Perceptual experience*, Oxford: Oxford University Press, pp. 481–503.
- Siegel, S., (2010a), Do visual experiences have contents? In Nanay, B., (Ed.), (2010), *Perceiving the World*, Oxford: Oxford University Press, pp. 333–68.
- Siegel, S., (2010b), *The Contents of Visual Experience*, New York: Oxford University Press.
- Siegel, S., (2014), Affordances and the contents of perception, In Brogaard, B., (Ed.), (2014), Does Perception have Content? New York: Oxford University Press, pp. 51-75.
- Silva J. F., and Yrjönsuuri M., (Eds.), (2014) *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland.
- Silva, J. F., (2014a), Augustine on Active Perception, In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 79-98.
- Silva, J. F., (2014b), Medieval Theories of Active Perception: An Overview, In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 117-146.
- Sincich, L. C., Park K. F., Wohlgemuth M. J., and Horton J. C., (2004), Bypassing V1: a direct geniculate input to area MT, *Nature Neuroscience*, 7, pp. 1123–1128,

doi: 10.1038/nn1318.

- Singh-Curry, V., and Husain M., (2009), The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy, *Neuropsychologia*, 47 (6), 1434-1448, doi: 10.1016/j.neuropsychologia.2008.11.033.
- Singhal A., Monaco S., Kaufman L. D., and Culham J. C., (2013), Human fMRI Reveals That Delayed Action Re-Recruits Visual Perception, *PLoS One*, 8 (9), doi: 10.1371/journal.pone.0073629. eCollection 2013.
- Sinigaglia, C., (2013), What type of action understanding is subserved by mirror neurons? *Neuroscience Letters*, doi: 10.1016/j.neulet.2012.10.016.
- Sinigaglia, C., and Butterfill S., (2015a), Motor representation in goal ascription. In Fischer M.H. and Coello Y. (Eds.) *Conceptual and Interactive Embodiment. Foundations of Embodied Cognition*, Vol. II, Routledge: Oxford UK, pp. 149-164.
- Sinigaglia, C., and Butterfill S., (2015b), A puzzle about the relations between thought, experience, and the motoric, *Synthese*, 192 (6), pp. 1923-1936.
- Smeets J. B. J., Brenner E., de Grave D.D. J., Cuijpers R.H., (2002), Illusions in action: consequences of inconsistent processing of spatial attributes, *Experimental Brain Research*, 147 (2), pp. 135–144, doi 10.1007/s00221-002-1185-7.
- Smith, P., (2013), Cézanne's “Primitive” Perspective, or the “View from Everywhere”, *The Art Bulletin*: 95 (1), pp. 102-119.
- Solomon, R., (Ed.), (2003), *Thinking About Feeling: Contemporary Philosophers on Emotions*, New York: Oxford University Press.
- Solomon, R., (2004), *In Defense of Sentimentality*, New York: Oxford University Press.
- Sperandio, I., Savazzi S., and Marzi C. A., (2009), Is simple reaction time affected by visual illusions? *Experimental Brain Research*, 201 (2), pp. 345-350, doi 10.1007/s00221-009-2023-y.
- Srivastava S., Orban G. A., De Mazière P. A., Janssen P., (2009), A distinct represen-tation of three-dimensional shape in macaque anterior intraparietal area: fast, metric, and coarse, *Journal of Neuroscience*, 29 (34), 10613–10626, doi: 10.1523/JNEUROSCI.6016-08.2009.
- Stein, B. E., Stanford T. R., Wallace M. T., Vaughan J. W., and Jiang W., (2004), Crossmodal spatial interactions in subcortical and cortical circuits. In Spence, C., and Driver, J. (Eds.), (2004), *Crossmodal Space and Crossmodal Attention*, Oxford: Oxford University Press, pp. 25–50.
- Taira, M., Nose I., Inoue K., and Tsutsui K., (2001), Cortical areas related to attention to 3D surface, *Neuroimage,* 14 (5), pp. 959-966.
- Tappolet, C., (2003), Emotions and the Inelligibility of Akratic Action, In Stroud S., and Tappolet C., (Eds.) (2003), *Weakness of Will and Practical Irrationality*, Oxford: Clarendon Press, pp. 97–120.
- ter Horst, A. C., van Lier, R, Steenbergen, B., (2011), Spatial dependency of action simulation, *Experimental Brain Research*, 212, pp. 635–644 doi: 10.1007/s00221-011-2748-2.
- Theys T., Srivastava S., van Loon J., Goffin J., Janssen P., (2012a), Selectivity for three-dimensional contours and surfaces in the anterior intraparietal area. *Journal of Neurophysiology,* 107 (3), pp. 995–1008.
- Theys T., Pani P., van Loon J., Goffin J., and Janssen P., (2012b), Selectivity for three-dimensional shape and grasping-related activity in the macaque ventral premotor cortex, *Journal of Neuroscience*, 32 (35), pp. 12038–12050.
- Theys T., Pani P., van Loon J., Goffin J., Janssen P., (2013), Three-dimensional shape coding in grasping circuits: a comparison between the Anterior

Intraparietal Area and Ventral Premotor Area F5a, *Journal of Cognitive Neuroscience*, 25 (3), pp. 352–364.

- Theys, T., Romero M. C., van Loon J., and Janssen P., (2015), Shape representations in the primate dorsal visual stream, *Frontiers in Computational Neuroscience*, 9 (43), doi: 10.3389/fncom.2015.00043.
- Thill, S., Caligiore D., Borghi A. M., Ziemke T., and Baldassarre G., (2013), Theories and computational models of affordance and mirror systems: An integrative review, *Neuroscience and Biobehavioral Reviews,* 37 (3), pp. 491–521, doi: 10.1016/j. neubiorev.2013.01.012.
- Triberti, S., Repetto C., Costantini M., Riva G., and Sinigaglia C., (2015), Press to grasp: how action dynamics shape object categorization, *Experimental Brain Research*, 234 (3), pp. 799-806, doi: 10.1007/s00221-015-4446-y.
- Tsutsui, K., Sakata H., Naganuma T., and Taira M., (2002), Neural correlates for perception of 3D surface orientation from texture gradient, *Science*, 298 (5592), pp. 409-412.
- Tsutsui, K., Taira M., Sakata H., (2005), Neural mechanisms of three-dimensional vision, *Neuroscience Research*, 51 (3), pp. 221–229.
- Tucker, M., and Ellis R., (1998), On the relations between seen objects and components of potential actions, *Journal of Experimental Psychology: Human Perception and Performance*, 24 (3), pp. 830–846.
- Tucker, M., and Ellis R., (2001), The potentiation of grasp types during visual object categorization, *Visual Cognition*, 8 (6), pp. 769–800.
- Tucker, M., and Ellis R., (2004), Action priming by briefly presented objects, *Acta Psychologica*, 116 (2), pp. 185–203, doi: 10.1016/j.actpsy.2004.01 .004.
- Tuominem, M., (2014), On Activity and Passivity in Perception: Aristotle, Philoponus, and Pseudo-Simplicius, In Silva J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 55-78.
- Turella, L., and Lignau A., (2014), Neural correlates of grasping, *Frontiers in human neuroscience*, 8 (686), doi: 10.3389/fnhum.2014. 00686.
- Turnbull, O. H., Driver J., and McCarthy R. A., (2004), 2D but not 3D: pictorial depth deficits in a case of visual agnosia, *Cortex*, 40 (4-5), pp. 723–738.
- Vainio, L., Hammaréén L., Hausen M., Rekolainen E., and Riskilää S., (2011), Motor inhibition associated with the affordance of briefly displayed objects, *Quarterly Journal of Experimental Psychology*, 64 (6), pp. 1094–1110.
- van Elk M., van Schie H. T., Neggers S. F., and Bekkering H., (2010), Neural and temporal dynamics underlying visual selection for action, *Journal of Neurophysiology,* 104 (2), pp. 972–983.
- van Elk, M., Paulus M., Pfeiffer C., van Schie H. T., and Bekkering H., (2011), Learning to use novel objects: a training study on the acquisition of novel action representations, *Consciousness and Cognition,* 20 (4), pp. 1304–1314, doi: 10.1016/j.concog.2011. 03.014.
- van Elk, M., van Schie H. T., and Bekkering H., (2014), Action semantics: A unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge, *Physics of Life Reviews,* 11, pp. 220–250, doi: 10.1016/j. plrev.2013.11.005.
- Varela F., Thomson E., Rosh E., (1991), *The Embodied Mind*, The MIT Press, Cambridge (MA).
- Vargas M., and Yaffe G., (2014), *Rational and Social Agency. The Philosophy of Michael*

Bratman, Oxford University Press.

- Verhoef, B. E., Vogels R., Janssen P., (2010), Contribution of inferior temporal and posterior parietal activity to three-dimensional shape perception, *Current Biology*, 20 (10), pp 909–913.
- Viljanen, V., (2014), Spinoza on Activity in Sense Perception. In In Silva, J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 241-254.
- Vishton, P., and Fabre E., (2003), Effects of the Ebbinghaus illusion on different behaviors, *Spatial Vision*, 16, pp. 377–92.
- Vishwanath, D., (2007), Is Focal Blur an Absolute Cue to Depth? *Journal of Vision,* 7 (9): pp. 845, 845.
- Vishwanath, D., (2008), The Focal Blur Gradient Affects Perceived Absolute Distance, *Perception* 37, 130.
- Vishwanath, D., (2010), Visual Information in Surface and Depth Perception: Reconciling Pictures and Reality, in Albertazzi L., van Tonder G. J., and Vishwanath D., (Eds.), (2010), *Perception beyond Inference, the Information content of visual processes*, Cambridge, MA: MIT Press, pp. 201 -240.
- Vishwanath, D., (2014), Toward a New Theory of Stereopsis, *Psychological Review*, 121 (2), pp. 151–178, doi: 10.1037/a0035233.
- Vishwanath, D., Girshick A. R., and Banks M. S., (2005), Why pictures look right when viewed from the wrong place, *Nature Neuroscience*, 8 (10), pp. 1401–1410.
- Vishwanath, D., and Hibbard P., (2010), Quality in depth perception: The plastic effect, *Journal of Vision*, 10 (42), doi: 10.1167/10.7.42.
- Vishwanath, D., and Hibbard P., (2013), Seeing in 3D with just one eye: Stereopsis in the absence of binocular disparities, *Psychological Science*, 24 (9), pp 1673–1685, doi: 10.1177/0956797613477867.
- Voltolini, A., (2013), Why, as responsible for figurativity, seeing-in can only be inflected seeing-in, *Phenomenology and the Cognitive Sciences*, 14 (3), pp. 651-667, doi: 10.1007/s11097-013-9335-x.
- Watt, S. J., and Bradshaw M.F., (2003), The visual control of reaching and grasping: binocular disparity and motion parallax, *Journal of experimental psychology. Human perception and performance*, 29 (2), pp. 404-415.
- Wee, C., (2014), Descartes and Active Perception, In Silva, J. F., and Yrjönsuuri M., (Eds.), (2014), *Active Perception in the History of Philosophy from Plato to Modern Philosophy*, Springer, Switzerland, pp. 207-222.
- Westwood, D., Danckert J., Servos P., and Goodale M., (2002), Grasping two-dimensional images and three-dimensional objects in visual-form agnosia, *Experimental Brain Research*, 144 (2), pp. 262- 267.
- Wollheim, R., (1987), *Painting as an art*, Princeton University Press: Princeton.
- Wollheim, R., (1998), On pictorial representation, *Journal of Aesthetics and Art Criticism*, 56, pp. 217–226.
- Wollheim, R., (1999), *On the Emotions*, New Haven: Yale University Press.
- Wollheim, R., (2003), What makes representational painting truly visual? *Aristotelian Society Supplementary* 77, pp. 131-147.
- Wong, H. Y., (2009), On the Necessity of Bodily Awareness for Bodily Action, *Psyche*, *Consciousness in Natural and Cultural Contexts* (CNCC), Special Issue, 15.
- Wong, H. Y., (2010), Bodily Awareness and Bodily Agency, in O'Connor T., and Sandis C., (2010) (Eds.), *A Companion to the Philosophy of Action*, Blackwell.
- Wong, H. Y., (2015), On the Significance of Bodily Awareness for Action,

Philosophical Quarterly, doi: 10.1093/pq/pqv007.

- Young, G., (2006), Are different affordances subserved by different neural pathways? *Brain and Cognition*, 62 (2), pp. 134–142.
- Rossetti, Y., Pisella L., Vighetto A., (2003), Optic ataxia revisited: Visually guided action versus immediate visuomotor control, *Experimental Brain Research,* 153 (2), pp. 171–179, doi: 10.1007/s00221-003-1590-6.
- Zhu, J., and Thagard P., (2002), Emotion and action, *Philosophical Psychology*, 15 (1), pp. 19-36.
- Zipoli Caiani, S., (2013), Extending the notion of affordance, *Phenomenology and the Cognitive Sciences*, 13, pp. 275-293, doi: 10.1007/s11097-013-9295-1.